A3525 no. 148

LIBRARY FILE COPY
INTERMOUNTAIN STATION

United States Department of Agriculture

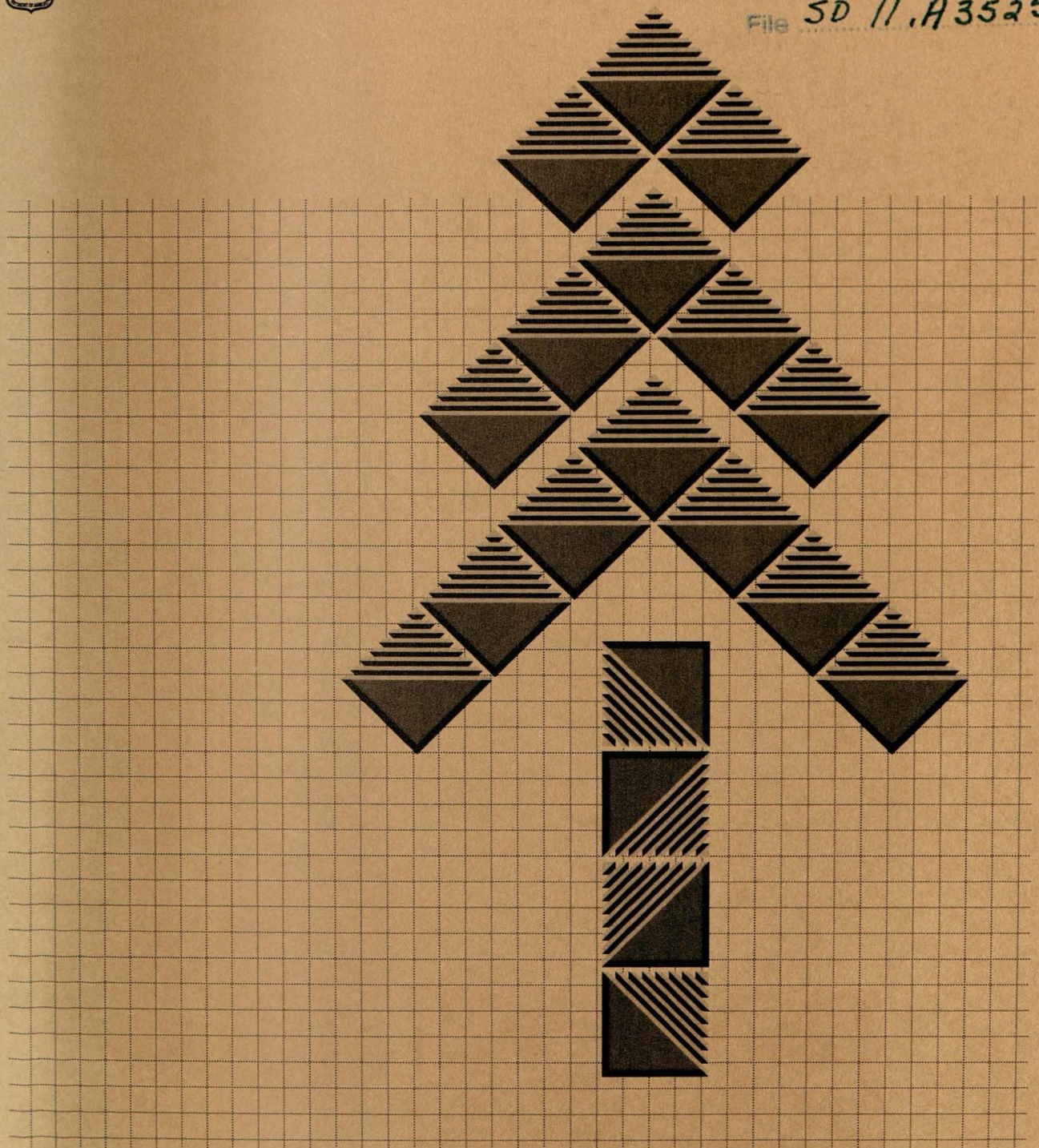
Forest Service

North Central Forest Experiment Station

General Technical Report **NC-148**

Quantity and Quality in Forest Research

FS INFO–Intermountain
MONTHLY ALERT
Edition ... NOV 1987
Item No. 42
File SD 11.A3525 no.148



Preface

The theme selected for the S6.09 meeting corresponding with the XVX IUFRO World Congress was quantity and quality in forest research. In particular I wished the last meeting of S6.09 under the leadership of Roelof Oldeman and myself to be on quality in forest research. As Robert Lee reminds the reader, improving the quality of forest research was one of the important reasons for 6.09 coming into being in 1981. How well each author hit the mark will, of course, be decided by you, the reader.

When the concepts of quantity and quality are applied to the issue of scientific research, two immediate issues are raised: meaning and measurement. What is quantity? What is quality? Also, how do we measure quantity? How do we measure quality? Clearly, the issue of quality is not a political or administrative one. A few objective criteria exist that have been developed over the years and are accepted by the wider scientific community. The purpose of papers at our meeting was to explore what those criteria might be in terms of each of our own specialities.

I thank the authors who presented talks in Montreal and, in some cases significantly expanded their verbal comments in written form for publication here. For a variety of reasons only two of the papers were published in the official proceedings of IUFRO Division 6. My contribution here is expanded over the version in Division 6 Proceedings. To signify this I have appended a 'II' to the title. A special thanks goes to my associate H. N. Phan who converted our manuscripts into TEX.

Rolfe A. Leary
St. Paul, Minnesota
USA

North Central Forest Experiment Station
Forest Service—U.S. Department of Agriculture
1992 Folwell Avenue
St. Paul, Minnesota 55108
Manuscript approved for publication September 26, 1991
1991

Quantity and Quality in Forest Research

Invited Papers Delivered
at the
XIX World Congress
of
International Union of Forest Research Organizations
Montreal, Canada

August 5 - 11, 1990

Papers compiled by
Rolfe A. Leary
Leader(1981–1990), Subject Group 6.09
Philosophy and Methods of Forest Research

USDA Forest Service
North Central Forest Experiment Station
1992 Folwell Avenue
St. Paul, Minnesota 55108
USA

Contents

Scholarship versus Technical Legitimation: Avoiding Politicization of Forest Science	Robert G. Lee	1
Scholarship in Forest Models	Richard G. Olderwald	12
The Identification and Analysis of Problems	Kenneth M. Brown	19
Quality of Modelling: Uniting Opposites	Fedor N. Semevsky and Boris Zeide	33
Cogency in Forest Research: II	Rolfe A. Leary	44

SCHOLARSHIP VERSUS TECHNICAL LEGITIMATION: AVOIDING POLITICIZATION OF FOREST SCIENCE

Robert G. Lee
College of Forest Resources
University of Washington
Seattle, Washington 98195
USA

INTRODUCTION

The 1981 IUFRO Congress in Kyoto, Japan was the beginning of the subject group in Philosophy and Methods of Forest Research. Two of the four objectives of this group were to improve the quality of forest research and to contribute to the maturing of forestry as a science (Leary 1981). Leary also suggested two criteria for judging the maturity of forest science: 1) development of a terminology that would reduce reliance on uncritical assumptions and mutual understanding, and 2) mathematization. I suggest that the necessity to adopt an explicit **logic of inquiry** be added as a third criterion. The applied, policy-oriented nature of much forestry research necessitates the adoption of more explicit rules for drawing inferences from observations. Development of terminology and mathematization do not meet this need, even though they can contribute significantly to the quality of empirical research.

This paper focuses on the distinction between scholarly and technical research to demonstrate why an explicit logic of inquiry is needed. Technical research seeks to provide means for predetermined ends. Emphasis is placed on justifying a choice among alternative means. Technical research focuses on discovering the best methods with which to support policy decisions, not on the policy decisions themselves.

One of the most striking examples of purely technical research was the study of wildfire problems in the United States. Ashley Schiff (1958) documented how the U.S. Forest Service designed and executed research projects over almost 50 years to prove its claim that fire was a destructive agent in forests and should be eliminated. A national commitment to minimizing the number and size of fires persisted until almost 1970. Aside from some changes in the South during the 1940's, all of the agency's fire-related research was dedicated to providing support for a policy commitment to excluding fire from all forests.

Scholarly research involves the comprehensive study of the larger processes in which the techniques are embedded. Assumed policy decisions are often challenged as a result of research into the nature of social, biological, or physical processes.

Investigation into the beneficial functions of fire and other natural disturbances required an understanding of basic ecological processes. Results from the Forest Service's own research on ecological processes and silvicultural options (not fire research) ultimately undermined its commitment to excluding all fire from forests.

Technical rationality has been far more common than scholarly research in forestry. A primary reason for such technical emphasis is the organizational context of policy-oriented research. Most forestry research is designed to serve the needs of private and public land management organizations. Understanding of basic biological, physical, social, or economic processes is less important than knowing how information can be used to serve organizational objectives. Facts are selected (or even distorted) to serve the organization's purposes (Bella 1987, Schiff 1962, Lindblom 1990). Many forestry research organizations, including universities and public agency research groups, find themselves doing technical research funded by land management organizations. Opportunities for research of a more fundamental (scholarly) nature is more limited, and often requires independent funding.

Contemporary proposals for rapid change in forest management practices necessary to incorporate new ecological information make it imperative that we don't repeat the mistakes of the past by doing research to support new policy commitments. Development of a terminology and mathematization will not prevent researchers from pursuing research with the objective of justifying new practices. A mature approach to forest science requires the adoption of procedures that will prevent the selection or distortion of facts.

The need for an explicit logic of inquiry will be presented in three parts. It will begin with a discussion of an overall logic of inquiry. Standards for judging the maturity of forestry research incorporating rules of inference will then be established. Finally, competing interpretations of the Yellowstone fires of 1988 will be used to illustrate the need for a self-conscious logic of inquiry in forestry research.

A LOGIC OF INQUIRY

Research in forest science must satisfy three methodological requisites to be considered scholarly rather than technical. First, researchers must be committed to making predictions and providing explanations rather than issuing or implementing prescriptions for management actions. To prescribe particular practices is to take on the role of citizen or public official. We step out of our role as scientists when we prescribe actions, since actions are always goal directed and, hence, involve uncritical commitments to values and beliefs. We forfeit our role as scientists when we advocate either fire exclusion or allowing fires to burn as they would under "natural" conditions. Scientists are most scholarly when they explain cause and effect processes, define management alternatives, or predict the consequences of alternative

management actions, and leave prescription to citizen actions and public officials.

Researchers serve as research technicians when they seek alternative means for implementing an organizational purpose without studying the larger processes — especially when they avoid the gathering or reporting of facts that would challenge policy commitments. Even though technical researchers may not overtly advocate particular goals, the fact that their work is governed by a limited set of goals implies an advocacy stance toward their work.

To predict or explain, scientists must subject theories to empirical tests. Hence, the second methodological requirement for an explicit logic of inquiry is that theories should permit falsification. Theories are best tested by deriving refutable implications from them, and then collecting data that could be inconsistent with those implications. Refutable implications take the form of hypotheses or propositions that are susceptible to challenges by empirical data. Researchers seeking to support a given policy will look for facts that are consistent with the organization's policy commitments. U.S. Forest Service researchers selected facts showing that fire was a destructive process, and discounted facts showing that fire was necessary for the regeneration of many plant species, including valuable coniferous timber species (Schiff 1962).

The third methodological requirement is the adoption of strong inference (Platt 1964). Implementation of strong inference requires: (1) formulating alternative hypotheses; (2) devising an experiment (or quasi-experiment or simulation) with alternative possible outcomes, each of which excludes one or more hypotheses; and (3) rigorously carrying out the experiment (quasi-experiment or simulation).

Strong inference provides an appropriate standard for guiding the maturation of inductive forestry research. Its logic is especially compelling. The best analogy is the act of shaping a tree by pruning it. Each fork provides an opportunity for growth to continue in a given direction. Alternative paths for future growth are eliminated by tests that falsify competing hypotheses. Sequential development of surviving hypotheses and subhypotheses give shape to theory, just as surviving branches give shape to the tree. Most scientists are involved in shaping a theory that explains a particular process or yields predictions about cause and effect relations implied by the process. Secondarily, they may challenge hypotheses derived from competing theories, or major stems or even separate trees (disciplines). Technical research involves limited opportunities for shaping a theory because it tends to focus on narrower objectives defined by preferred outcomes rather than on theoretical models of basic processes governing cause and effect relationships. The shape of the tree does not emerge along with a progression of studies, but is predetermined by policy commitments.

STAGES OF METHODOLOGICAL DEVELOPMENT

Methodological progress in developing theory may be assessed by utilizing as a standard of comparison the stages through which more mature theories have developed: ruling theory, working hypotheses, and multiple working hypotheses (Chamberlin 1897). These stages represent a progression from the defense of favored ideas to the utilization of strong inference. At each successive stage, empirical evidence assumes a greater role in the evaluation of theoretical statements.

Ruling theory is generally found at the earliest stage of methodological development. During their formative stage, theories are elaborated and extended to encompass many events and phenomena. Scientists take pride of authorship in the theory and initially seek facts that will support the application of theory to a wide variety of phenomena. The intellectual offspring becomes an object of considerable affection and is defended against competing ideas. Plausible explanations are adopted as theories and, when rigorously defended, become "ruling theories" (Also see Kuhn 1962).

Parental pride and protectionism leads investigators to ignore facts that are inconsistent with ruling theory, and when challenged, can lead them to actively repress facts that threaten their favored theories. When such repression becomes apparent to others and draws criticism, improvement is sought in the method of the working hypothesis. The purpose of the working hypothesis has been to challenge favored ideas with facts. Facts are assembled and their relations examined to challenge the hypothesis and to reshape the theory. Yet there is also a danger that working hypotheses may become an intellectual child to which investigators cling, even to the point of defending them with paternalistic passion.

To guard against this tendency, researchers may adopt the method of multiple working hypotheses. Multiple working hypotheses divide the paternalistic affections among several competing hypotheses. Parentage of a family of competing hypotheses promotes impartiality in the investigation of facts. At this point researchers begin practicing the methodology of strong inference. Within a given theory, crucial experiments involving competing hypotheses facilitate progressive exclusion of branches. Although much forestry research does not allow true experiments, quasi-experimentation, simulation, and other approximations of experimentation provide opportunities to practice the method of multiple working hypotheses.

The adoption of a methodology involving strong inference is neither a simple nor an unambiguous process. Many practical problems make hypothesis testing and evaluation of results very difficult. Skeptics may argue that a particular test is flawed because research procedures reflect inappropriate test conditions (measurement theory, simplifying assumptions, or selection of observations) rather than the truth value of a proposition. Even the most rigorously conducted research will be

subjected to such criticisms. But as a body of evidence emerges, the accuracy of the theory will become apparent.

These three stages of methodological development provide standards by which to assess the maturity of forestry research. Use of multiple, competing working hypotheses, coupled with the logic of strong inference, represent the most advanced stage of research. Platt (1964) argued that sciences at this stage in their development make exceptionally rapid progress in developing and elaborating theory. He cited biochemistry as a notable example. However, the sort of policy-oriented research common to forestry seldom reaches this stage. Research controlled by land management organizations may be closer to the stage of ruling theory.

Bella (1987, p 369) goes so far as to state that:

> Organizations tend to systematically distort information in self-serving ways. Such distortions do not depend... [on] deliberate falsifications by individuals. Instead, people who are competent, hard-working and honest can sustain systematic distortions by merely carrying out their organizational roles. Unchecked by outside influences or the undeniable realities of catastrophic failures, organizations can sustain self-serving distortions. The potential for catastrophic consequences is significant.

As a scientist, Szilard (1961, p. 42) makes a similar point about the openness of society to alternative ways of thinking when he states that Americans "... were free to say what they think, because they did not think what they were not free to say."

Technical research subjects itself to a far greater risk of self-deception, including selection of evidence favorable to sustaining organizational goals, ruling theories, or prerogatives. Scholarly research has a greater chance of identifying alternative approaches to land management because it seeks to challenge multiple, competing hypotheses that may suggest entirely different ways of thinking about problems and solutions—yielding alternative interpretations of how processes actually work and unconventional findings about the relationships between causes and effects. Scholarly research attempts to avoid self-deception or "group think" (See Lindblom's 1990 discussion of impairments to inquiry.) by cultivating doubt and adopting explicit rules for exercising formalized skepticism.

YELLOWSTONE FIRES OF 1988: FROM RULING THEORY TO A WORKING HYPOTHESIS

Recent analysis of the 1988 Yellowstone fires provides a convenient illustration of progress in forestry research. A ruling theory of natural regulation advanced by Romme and Despain (1989a and 1989b) has been challenged by a working hypothesis (Bonnicksen 1990). Theoretical elaboration has been suggested that could yield

a family of multiple working hypotheses subject to falsification (Christensen 1989, Romme 1990, and Turner 1990).

According to many critics, research at Yellowstone National Park has been insulated from external review and criticism by the scientific community (Robbins 1963; also see Chace 1988 for a restatement of this criticism). Park management policies at Yellowstone have relied on a theory of "natural regulation." Allowing "nature to take its own course" was defended as the most appropriate theory for sustaining animal populations, forests, and the integrity of natural ecosystems. The primary goal of the park has been to allow natural disturbances to maintain equilibrium in "an ecosystem shaped primarily by natural geological and ecological processes" (Houston 1971).

A "let burn" fire policy (later renamed a "prescribed natural fire policy") was adopted by the National Park Service in 1972. When implemented at Yellowstone, fires were allowed to take their "natural course" and were not suppressed. This approach to fire management continued until the unusual burning conditions of 1988 resulted in large uncontrolled fires. The Secretary of Interior ordered full suppression of several raging wildfires mid-way through the summer of 1988. A more restrictive interpretation of the national fire management policy was adopted by all federal land management agencies after reviews of the 1988 fire season had been completed.

Following the 1988 fires, researchers committed to the goal of natural regulation presented evidence to support their claim that the fires of 1988 were entirely natural and represented an approximately 200 to 300 year fire cycle in which mass conflagrations would burn over large areas of Yellowstone National Park (Romme and Despain 1989b). To support this interpretation, Romme and Despain presented data showing the proportion of the park that burned within fifty year periods ranging from 1690 to 1988 (See Figure 1). The data appear to support the claim that the large area burned in 1988 mimics the pattern of large fires that occurred almost 300 years earlier.

This claim was challenged by Bonnicksen (1990) when he regrouped Romme and Despain's fire history data into the ten year periods from which they had assembled the fifty year records. Bonnicksen's graph shows that the period of 1690 to 1740 had significant fires in every decade, while the period from 1940 to 1988 was nearly free of fires with the exception of 1988 (See Figure 2). These data show that the historic fires were smaller and less catastrophic than the fires of 1988. Hence, there is sufficient evidence for Bonnicksen to falsify the claim that the 1988 fires were "natural" because they resembled historical norms. At no point in the documented history of Yellowstone vegetation did such a large fraction of the park burn during one fire season, or even during less than three or four decades.

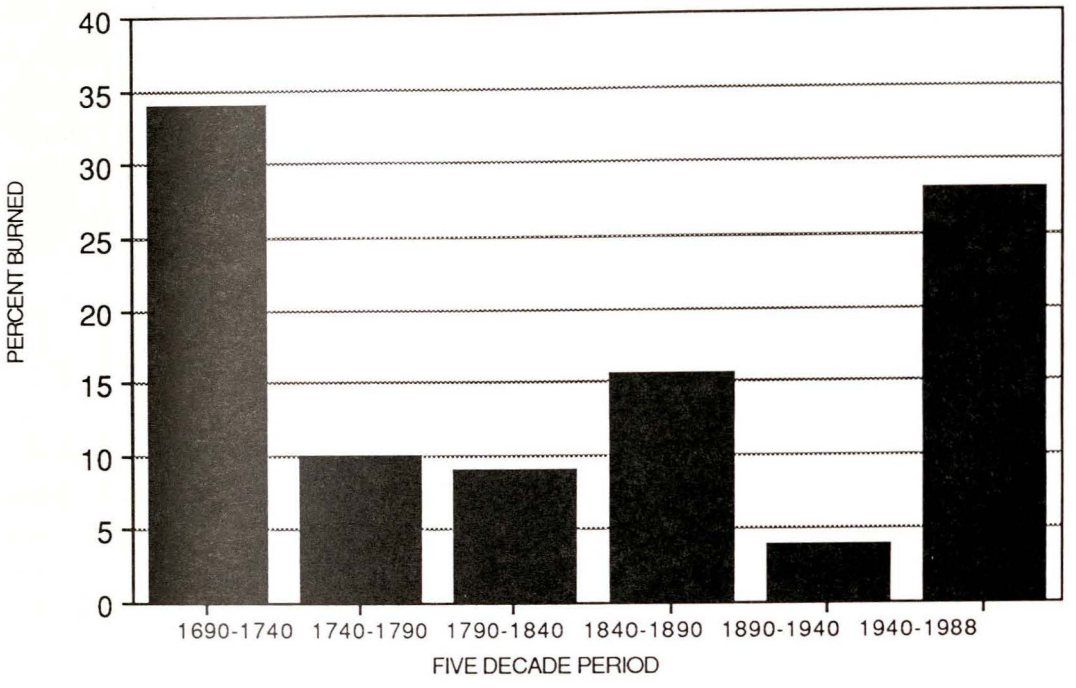


Figure 1. Proportion of Yellowstone National Park burned by six 5 decade periods (1690–1988). Data from Romme and Despain 1989b.

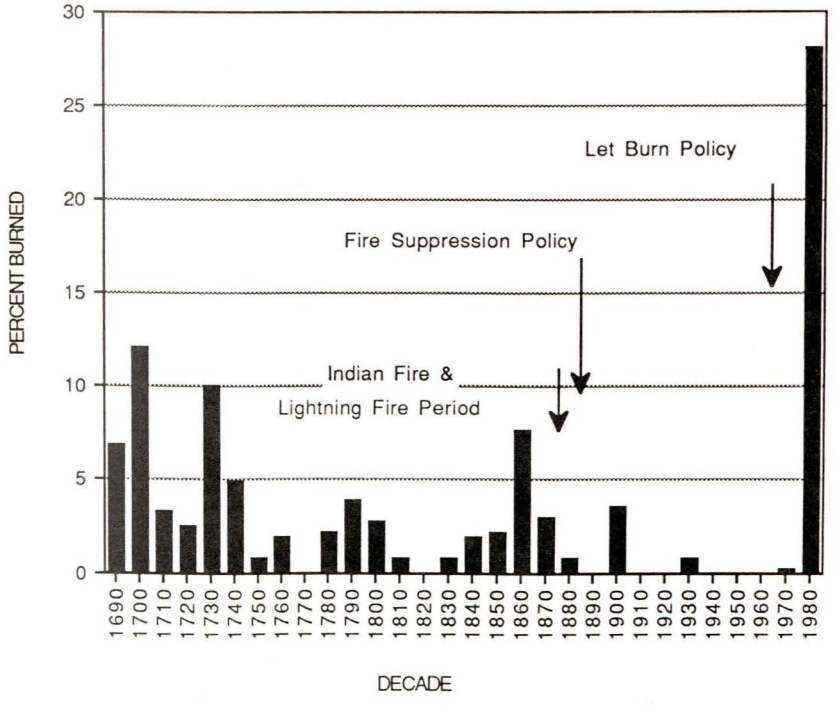


Figure 2. Proportion of Yellowstone National Park burned by decade (1690–1980). Data from Romme and Despain 1989b.

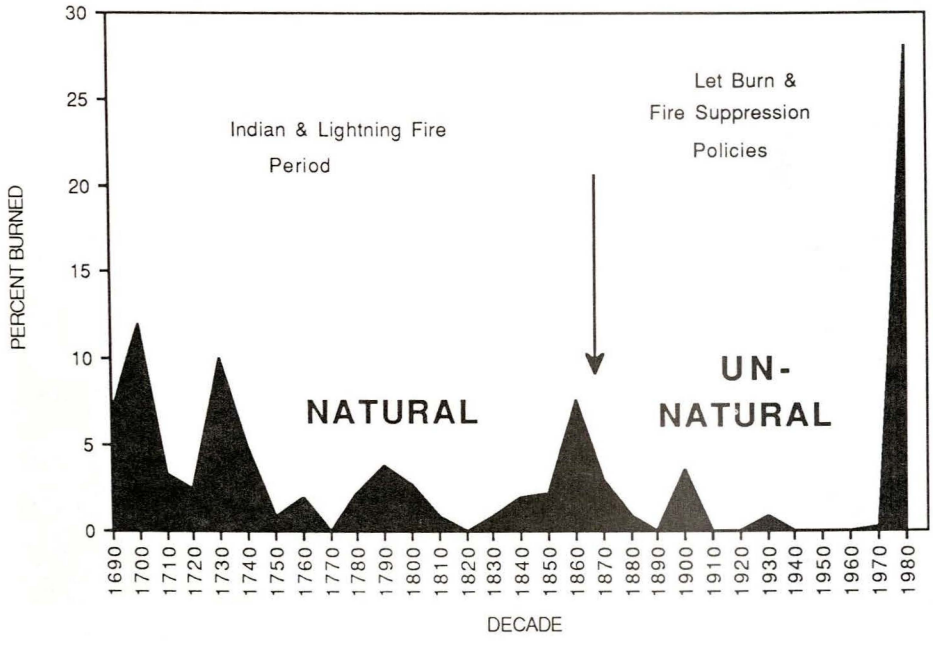


Figure 3. Proportion of Yellowstone National Park burned by decade (1690–1980). Data from Romme and Despain 1989b.

Bonnicksen also challenged Romme and Despain's claim that there were no significant effects on vegetation patterns from the implementation of the fire suppression policy in 1886 or removal of the Indians in 1878. He presented Romme and Despain's own data to show that "The long history of frequent fires ended almost abruptly between 1878 and 1886..." (See Figure 3). This evidence is sufficient to raise serious questions about Romme and Despain's claims.

However, Bonnicksen (1990, p. 28) went on to claim that "This means that available fuels did accumulate in Yellowstone for over a century. These fuels were critical to the size and severity of the 1988 wildfires." While plausible, Bonnicksen did not challenge this latter claim with sufficient evidence. Rival explanations, such as a sustained draught caused by early stages of global warming, have also been advanced to explain the unusual fires of 1988. Hence, at this point Bonnicksen began to reveal paternalism toward his own favored hypothesis.

Christensen (1989), previously a proponent for theories of natural regulation in national park management, has responded to the Yellowstone fires of 1988 with a call for theoretical elaboration. He assembled evidence that led him to question the assumption that equilibrium states in natural ecosystems will be achieved by allowing natural disturbances to occur. He postulated that the ecological processes affecting landscape patterns are not deterministic (explicitly rejecting the "Calvinist biology" of predestined equilibrium states arising from natural regulation), but instead

involve a complex interaction between stochastic and deterministic processes. Both Turner (1990) and Romme (1990), also independent, non-agency scientists, have recently suggested the possibility of non-equilibrium processes when they reported preliminary results from their ecological studies of post–1988 Yellowstone landscapes. Such scholarly consideration of basic ecological processes at the landscape scale suggests that research on natural ecosystem management may be entering the stage of multiple working hypotheses and strong inference.

CONCLUSIONS

As exhibited by the case of the 1988 Yellowstone fires, ruling theories can significantly distort forestry research. The technical training of foresters and other natural resource managers, coupled with an organizational context that demands "policy-relevant" research, have not been conducive to the maturation of forest science. Social and geographic isolation of agency researchers can further contribute to self-deception and "group think."

Forestry research is easily politicized because it is readily utilized by public and private land management organizations and environmental interest groups. Policy preferences and commitments can distort the way in which research is designed, conducted and interpreted. An explicit logic of inquiry is essential for guarding against such self-deception and "group think." The maturity of forest science cannot be judged by terminology and mathematization alone. The progression from ruling theory to multiple working hypotheses to strong inference, appears to meet the need for an additional criterion to evaluate the maturity of forest science.

LITERATURE CITED

Bella, D.A. 1987. "Organizations and Systematic Distortion of Information." *Journal of Professional Issues in Engineering* 113:360–370.

Bonnicksen, T.M. 1990. "Restoring Biodiversity in Park and Wilderness Areas: An Assessment of the Yellowstone Wildfires," Pp. 25–32 In Allen Rasmussen (ed), *Wilderness Areas: Their Impacts* Proceedings of a Symposium. Utah State University, Cooperative Extension Service, Logan, Utah.

Chace, A. 1988. *Playing God in Yellowstone.* New York, London: Harcourt, Brace, and Jovanovich.

Chamberlin, C. 1897. "The Method of Multiple Working Hypotheses." *Journal of Geology* 5:837.

Christensen, N. 1989. Ecosystems Northwest Seminar presented at the University of Washington, College of Forest Resources, November.

Houston, D.B. 1971. "Ecosystems in National Parks." Science 172:648–651.

Kuhn, T.S. 1962. *The Structure of Scientific Revolutions.* Chicago: The University of Chicago Press.

Leary, R.A. 1981. "Prospectus for a New Subject Group: Philosophy and Methods of Forest Research." Paper presented at Division 6, XVII IUFRO World Congress, Kyoto, Japan.

Robbins, W.R. 1981. A Report by the Advisory Committee to the National Park Service on Research of the National Academy of Sciences-National Research Council. Washington, D.C.: National Academy of Sciences-National Research Council.

Romme, W.H. 1990. Ecosystems Northwest Seminar presented at the University of Washington, College of Forest Resources, October.

Romme, W.H. and D.G. Despain. 1989a. "The Long History of Fire in the Greater Yellowstone Ecosystem." *Western Wildlands* 15 (2):10–17.

Romme, W.H. and D.G. Despain. 1989b. "Historical Perspectives on the Yellowstone Fires of 1988." *BioScience* 39 (10):695–699.

Schiff, A. 1962. *Fire and Water: Scientific Heresy in the Forest Service.* Cambridge, Mass: Harvard University Press.

Szilard, L. 1961. *The Voice of the Turtle.* New York: Simon and Shuster. (Quoted in Lindblom (1990), p. 63)

Turner, M. 1990. Ecosystems Northwest Seminar presented at the University of Washington, College of Forest Resources, October.

SIMPLICITY IN FOREST MODELS

Richard G. Oderwald
Forestry Department
Virginia Polytechnic Institute and State University
Blacksburg, Virginia 24061–0324
USA

SUMMARY

Four definitions of model simplicity are presented and evaluated on ease of use, applicability, and reliance on observations. Three of the four definitions are based on observations as applied to models, rather than on the models themselves. The definition that does not depend on observations to establish simplicity is chosen as the most useful.

Keywords: simplicity, models

INTRODUCTION

According to Egon Spengler,[1] "...it's just that in science we always look for the simplest explanation." But what is "simplest", and can we use it once we've got it?

The intent of this paper is to present some alternative definitions of simplicity in models, describe the characteristics of each alternative, and to point out which definition I believe is the most useful in forestry models. The alternative definitions presented here are by no means inclusive. Many other definitions of simplicity have been and can be proposed. I have included those definitions I have heard expressed or felt were being used, even if they were not expressed, and which I consider to have some chance of successful application.

I have based many of the opinions expressed here on the works of Karl Popper, The Logic of Scientific Discovery in particular, an excellent paper by G. R. Dolby in "Biometrics", The Semantic Conception of Theories and Scientific Realism by Frederick Suppe, and The Structure of Scientific Revolutions by T. S. Kuhn. A bibliography of these references appears at the end of this paper, but I have not ascribed particular ideas to particular sources.

[1] "Ghostbusters II", Columbia Pictures Industries, Inc., 1989.

DEFINITIONS OF SIMPLICITY

Simplicity as Least Degree

In this definition the model with the lowest mathematical degree is the simplest. As an example consider the simplest curve through a set of X, Y observations. The simplest curve is that curve with the least departure from linearity or the curve with the lowest degree that explains these observations. In this definition the model of a straight line, $Y = aX$, is simpler than the model of a quadratic, $Y = aX^2$ which is simpler than a cubic, $Y = aX^3$.

This definition is founded on the principle that generalizations can be achieved by induction from observations. Some group of observations are made as the result of experiment or measurement. These observations are examined for possible explanations for their behavior, and, on the basis of this examination, an explanation, a model, is developed. The simpler, as defined above, this model is, the better it is taken to be.

There are three particular problems with this definition. First there is no advantage in this definition of simplicity. One model is chosen over another because the first is simpler, but there it ends. No further use can be made of the simplicity. In the example above, all models (equations) are of the form $Y = aX^b$. Having a power, b, equal one serves only to save ink or paper. It is not easier to arrive at the hypothesized model, nor to justify that arrival.

Second, defining simplicity in this way is easy for models with few variables, parameters, or equations, but determining which of two distance-dependent, individual tree growth models is simpler by these lights is very difficult or plain impossible. Would "simplest" be determined by comparing equations term by term and equation by equation, possible only when both models have terms and equations that can be matched? If the models were comparable to this extent, they would be so similar that "simplest" would have no meaning. Perhaps counting the number of equations in each model would establish "simplest". But rearrangement is always possible. An alternative for regression based models could be that the simplest model is the one with the fewest minimized linear sums of squares. This would be easy to determine but ignores the variables and parameters that go into each model, and would make all single equation models equally simple regardless of the number of variables or parameters. Thus, for models with many variables, parameters, and equations we have a double whammy; simplicity is difficult to determine and provides no advantage once it is determined.

The third problem with this definition of simplicity is that it is based on the observations used to develop the model, not on the model itself. The model is arrived at

by induction from particular observations, by "fitting" (not necessarily regression fitting) the model to the observations. The simplest model has the least degree or the fewest moving parts for this set of data. However, a set of new observations of the same phenomenon may call for another model that is judged to be simplest. This model is not necessarily the same as the previous model, even though the observations are of the same phenomenon. We then have two "simplest" models, one for each of the sets of observations of the same phenomenon. If we take the simplest of these two models as appropriate for both sets of observations, then at least one of the models was not simplest after all.

Simplicity as Fewest Parameters

A second definition of simplicity is that model with the fewest terms and/or parameters. For example, the model $Y = aX^b$ is simpler than $Y = aX^b + cY^d$ which in turn is simpler than $Y = aX^b + cY^d + eZ^f$. This definition is very much like the first definition and possibly could be included in it. I have separated it here since I have seen both definitions used explicitly and because there is a subtle difference between them. This difference is exemplified by the model $Y = aX^b$. In the first definition the value of b was a cause for differing simplicity. A value of one for b gave a model simpler than a value of two, which in turn gave a model simpler than a value of three, and so on. In this second definition of simplicity the presence of the parameter is the distinguishing characteristic; the actual value of b has no bearing on the degree of simplicity.

This is a more easily applied definition than the first one, particularly for the typical models used in forestry, since the estimated value of a parameter is not important. The model $Y = aX^{1.9}$ generally would not be considered to be simpler than $Y = aX^{2.1}$. The distinction between models with and without additional parameters is much easier to make. Also, in a regression based model inclusion of additional variables or parameters would change the value of parameters previously in the model, thereby invalidating any consideration of the value of any parameter as a criterion for simplicity.

However, this definition has the same problems as the first definition. No advantage beyond a savings in calculation time and printing space is bestowed by applying the definition. The one parameter model may be "simpler" than the two parameter model, but what of it?

This definition is also difficult to apply to models with many terms, but it is at least easier to apply than the first definition. We can count the number of parameters in the competing models and judge as simplest the one with the fewest parameters. Unfortunately, this implicitly assumes that the models are

hierarchical; each more complex model contain the next simplest model as a subset, as $Y = aX^b + cY^d$ contains $Y = aX^b$. Otherwise the definition will lead nowhere. For example, if two growth models both contain terms for age, site and density but one model takes density as number of trees per unit area, while the other takes density as basal area per unit area, both models have the same level of simplicity and yet are very different. We cannot use the definition of simplicity to separate and choose between these two models.

Finally, this definition, like the first definition, is based on the observations used to develop the model, not on the model itself. The definition pertains to the model with the fewest parameters that can be constructed to explain a particular set of observations. New observations on the same phenomenon will most probably result in a different "simplest" model that may have little in common with the previous "simplest" model. With this definition we arrive at the same contradiction we previously encountered.

Simplicity as Most Probable

A third definition of simplicity takes as simplest the model that is most probable or has the greatest likelihood of being a correct generalization of the observations. This probabilistic definition of simplicity is very often employed in model construction, although not as a measure of simplicity, through selection of models that have the greatest likelihood, as in maximum likelihood estimators. However, I do not mean to use the statistical principle of maximum likelihood as the definition of simplicity. The premise of model selection on this basis is that unless the description represented by the model is correct it is improbable or not very likely that the model would fit the observations so well, and that any other contending models have less likelihood.

As a definition of simplicity, rather than only a model selection technique, this avoids the problems of determining which model has the lowest degree of the parameters or which has the fewest number of parameters. It is also more objective in that it can be easily applied and agreed to by all who are dealing with the models. Often the very process used to estimate parameter values and select models carries within it the greatest likelihood tenet.

But this definition still depends on the observations used to develop the model, perhaps even more so than the first two definitions since comparison with observations (determining greatest likelihood) is the sole criterion of simplicity. Any change in the observations or use of a new set of observations changes the model chosen as simplest.

This difficulty is common to all three of the above definitions; simplicity in a model cannot stand alone but only as an adjunct to changeable observations. As is probably clear by now, I do not believe that any definition of simplicity that depends on the particular observations used to define a model is useful. Simplicity in a model must be a property of the model irrespective of the observations that may be brought to bear upon or within that model. Simplicity must be part and parcel of the model even if it is not possible to collect observations to support or refute the model.

Simplicity as Easiest to Disprove

The definition of model simplicity I propose (and the idea is certainly not original with me) is that the simplest model is the model which is easiest to disprove.

A common illustration of this definition is the model "All crows are black". The model is "simple" in that it can be easily disproved, at least in principle, by finding a white (or red or blue or green) crow. The simplicity of this model does not depend on any particular observations of crow colors, but only on the implication of the model that one or more non-black crows is sufficient to disprove the model. The model is unaffected by any number of observations of black crows, and once a single non-black crow is seen to disprove the model no number of continued observations of black crows can revive it.

Notice here that it is not necessary that the disproving tests be easy to make, only that they be easy to define. I could spend many years looking for non-black crows; the test is defined and can disprove the model, and that alone is sufficient for simplicity.

This definition has two advantages over those previously described. First, model simplicity is independent of observations that may (or may not) be made on the phenomenon the model represents. For example, the simplicity of a growth and yield model for a particular species of tree does not depend on which group of trees are selected for measurement or from which side of the mill the trees are taken. Also, model simplicity is maintained across different sets of observations. A model of economic behavior would have the same simplicity for observations made on the United States economy or on the economy of Mongolia. In one case or the other it may be possible to more easily collect information which would disprove the model, but the model simplicity does not change. The simplicity of the model resides in the model itself; observations have no role in establishing that simplicity.

Second, this definition of simplicity is useful in that it directs attempts to disprove the model. In the example of the crows, the statement of the model shows at the

same time the degree of simplicity and the means to disprove the model. Determining the simplicity of a tree growth model shows what tests can cause the model to be discarded and what observations may be needed to make these tests. This characteristic makes simplicity useful and worthwhile to determine.

However, simplicity by this definition may not be easy to apply. Simplicity for the model "all crows are black" can be determined, but models with more variables and parameters would seem to be another case entirely. On one hand, the more terms and parameters a model has the more opportunities there would seem to be to disprove it and the simpler the model would be. On the other hand, it may be difficult to decide whether the entire model or only one feature has been rejected. In general, the balder the model assertions, the more easily it can be disproved and the simpler it is. In this regard this definition may claim the second definition as a subset, since fewer parameters means more general claims and more simplicity.

Unfortunately, our models necessarily have many parameters (e.g. an individual tree, distance-dependent growth model) and determining simplicity and then using the result will be difficult. But the notion of simplicity as most susceptible to disproof still is advantageous. We can begin to examine our models in the light of simplicity and move toward a core of models that are less often disproved and that may have wide generality.

CONCLUSION

Model simplicity has been defined as least degree, fewest parameters, most probable, and easiest to disprove. My main objection to the first three definitions is their reliance on observations to establish simplicity, although there are considerations of ease of use and applicability of results as well. I suggest the use of model simplicity as easiest to disprove because it is independent of observations, even though application may be difficult.

LITERATURE CITED

Dolby, C.R. 1982. The role of statistics in the methodology of the life sciences. Biometrics, vol. 38, pp. 1069–1083.

Kuhn, T.S. 1970. The Structure of Scientific Revolutions, 2nd Ed., The University of Chicago Press: Chicago, 210 p.

Popper, K.R. 1968. The Logic of Scientific Discovery, Harper and Row, Publisher, Incorporated, New York, 480 p.

Suppe, F. 1989. The Semantic Conception of Theories and Scientific Realism, University of Illinois Press, Urbana and Chicago, 475 p.

Prepared December 15, 1989

THE IDENTIFICATION AND ANALYSIS OF PROBLEMS

Kenneth M. Brown
Lakehead University
Thunder Bay, Ontario, Canada P7B 5E1

SUMMARY

Beginning graduate students often have trouble finding and solving their thesis research problems. Since noticing problems and solving them is procedurally the same in research as in ordinary life, we might expect new graduate students to be good at doing research. But, too often, this is not the case. In this paper, I ask, "Why not?" and "What can faculty mentors do to help?"

I attempt to answer these questions by examining the "common sense" methods that are used to solve problems in ordinary life. I argue that common sense problem solving is actually quite complex. It seems easy only because we command so much factual and procedural knowledge about ordinary life situations. Using an ordinary problem from my own life as a case study, I try to show why research is often so difficult for new graduate students and what faculty mentors might do to help their students succeed.

Keywords: common sense, frame, novice vs. expert, problem solving, protocol, research, strateic plans, tactical, plans

INTRODUCTION

Noticing problems and solving them is procedurally the same in research as in ordinary life. So, you might suppose that beginning graduate students, who are often good at noticing and solving the problems of ordinary life, would step easily into the research role — and some do. But for many new graduate students, thesis research presents many difficulties.

What makes finding and solving research problems difficult for beginners? How do experts do it? What can faculty mentors do to help students get off to a good start? One way to learn about these topics is to study the "common sense" methods that we use to solve the problems of ordinary life. Perhaps when we understand what we do in the domain of ordinary life, we will see what it takes to extend those methods into the domain of research.

THE BOOK PROBLEM

During the next few minutes, I plan to tell you about one of my own recent experiences with problem solving in an ordinary life situation. In the first telling, I will recount the events as they happened without much attention to process. Then, I will reflect on the same episode, and, in the second telling, I will draw out the kinds of factual and procedural knowledge that made my "common sense" problem solving possible. Finally, I will draw a few conclusions regarding graduate students and their research.

Recognizing That A Problem Exists

Last fall I placed a classified advertisement for an out-of-print book in a trade newsletter called The Library Bookseller. My want ad was seen by an antiquarian bookseller in the eastern United States, and he offered to sell me the book for $75. The book was worth that much to me, and so I bought a money order and mailed it to the bookseller.

Several weeks went by and still the book had not arrived. Finally, I decided to call the bookseller, but to my dismay I found that I had forgotten his name and his address. Clearly, I had a problem. But, what could to do about it?

Analyzing the Problem

At this early stage, I framed the problem quite broadly (Fig 1). I had paid for a book that I had not received. This conception of the problem was too vague to act upon, however. So, I decided to try to discover what, if anything, had gone wrong, and then, if necessary, to try to do something to set things right.

What went wrong?

For purposes of discovering what went wrong, I realized that the entire transaction, from sending the money order to receiving the book, could be subdivided into 3 distinct segments:

- transmission of my money order to the book seller (This segment of the transaction involved the postal systems of two countries and perhaps the US Customs Service.)
- processing the purchase by the bookseller (This segment involved the bookseller receiving the cheque and shipping the book.)
- transmission of the book from the bookseller to my office mailbox (This segment involved the postal systems of two countries, Canada Customs, and the Shipping and Receiving Department of my own University.)

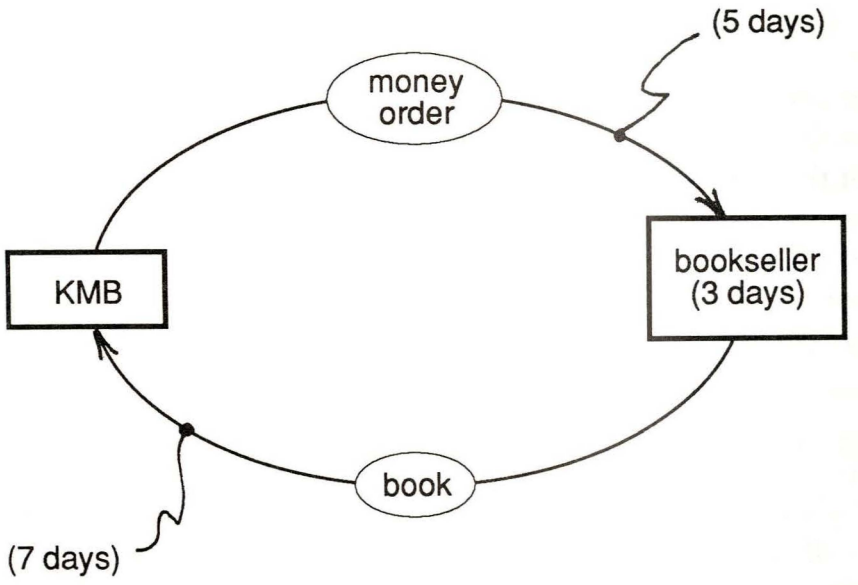


Fig 1. My initial, low resolution, model of the mail-order book transaction. Anticipated delays are given in parentheses.

This view of the transaction suggested a rough plan of action. First, I would call Shipping and Receiving and ask whether they had received the book. If not, I would set out to communicate with the bookseller so that I could ask him: Did you receive my money order? And if so - Did you ship the book? I decided to by-pass the postal and customs systems until I was sure that the book was in their hands.

I quickly determined that the Shipping and Receiving had not seen my book, and so my problem became one of trying to communicate with the bookseller.

Who is the bookseller?

In order to communicate with the bookseller, I had to track him down. Fortunately, I found a receipt for the money order, and on it the bookseller's business name, Elliot's Books. I still didn't know where the business was located except that it was somewhere in the eastern United States.

Then I remembered that Mr. Elliot had found me through The Library Bookseller. Perhaps I could find him the same way. I looked in an old issue of the Bookseller, and found the address of Elliot's Books: PO Box 6, Northford, CT.

At this point, I could have written to Mr. Elliot, but I was impatient, and so I

decided to try to speak to him on the telephone. I called directory assistance for Northford, CT, but they had no listing for Elliot's Books. Nor did they list a private residence under the surname "Elliot." The directory assistance operator volunteered that there were 3 Elliot's in a neighboring town, but I decided not to pursue that lead just yet.

Then I had an idea that quickly solved my problem. I called Northford directory assistance a second time and asked for the telephone number of the local Post Office. I called the Post Office and spoke to a nice lady who knew that Elliot's Books belonged to Mr. Elliot Ephraim. She was so nice, in fact, that she looked up his telephone number for me in her own directory. A few minutes later I was speaking to Mr. Ephraim.

In case you are curious, it seems that Mr. Ephraim delayed shipping the book because he was having second thoughts about his asking price - he thought it was too high. Two weeks later I had my book along with a $25 refund.

REFLECTING ON THE BOOK PROBLEM

Perception of the Problem

A problem exists whenever there is a gap between the present state (what is) and either

- a goal state (what is desired), or
- an expected state (what ought to be under present theory)

Disappointment and surprise are the harbingers of trouble.

The magnitude of a problem depends on two factors. The problem is *important* to the extent that we regard the gap between what is and what ought to be important. The problem is *difficult* if we find the gap hard to bridge.

Notice that many problems cannot be categorically classified as either important or difficult. These qualities depend on your point-of-view and experience. Whether or not a problem is important *to you* depends upon your point-of-view and appreciative system. Whether or not a problem is difficult *for you* depends upon your background knowledge of the situation and general problem solving skill.

One aspect of point-of-view is especially significant. It has to do with the distinction between *insiders* (those who have solid, first-hand experience with the problem domain) and *outsiders*. In many respects, problem solvers who are "inside" the problematic situation have a great advantage over those who are "outside." More often than not, insiders are in the best position to perceive problems, and they are best equipped with the factual and procedural information to solve them.

The exceptions to this rule are worth noting. Insiders can become blind to situations that an outsider *with the right background* might detect. Also, insiders may be so accustomed to standard ways of doing things that they are blocked from seeing creative solutions that *some* outsiders might see. But, the outsider cannot be just anyone. Outsiders who are capable of valuable insight are those who have a solid background in an analogous problem domain. It is debatable, therefore, whether an "outsider" with expert knowledge of an analogous system is *really* an outsider.

The inside view

It was easy for me to perceive the book problem, at least at a superficial level, because

- I had done something (ordered a book)

- I had an expectation (to receive the book within a reasonable period of time) which was based on a personal theory of how the system worked

- my expectation was not fulfilled (I did not receive the book within a reasonable period of time)

Furthermore, I perceived the book *situation* as a *problem* because of my personal sense of values. Not only did I sense a gap between what is and what is desired, but I judged the gap to be important. Why? First, because my $75 was important to me. Second, I placed a high value on owning that particular old book.

It is worth noting that there were several other participants in the book transaction: the clerk at the bank who sold the money order, post office employees who handled the envelope containing the money order, and the bookseller himself. But none of these players sensed the problem *as I did.*

From the points-of-view of those at the bank and the post office, there was no problem. But then, these individuals were all outsiders with respect to the transaction itself. From the point-of-view of Mr. Ephraim (the only other insider to the transaction), there was a problem, but it was not *the* problem that I sensed.

The Problem Analysis

Once I was committed to recovering either my book or my money, I had to figure out how to do it. I have already said *what I did.* Now I want to reflect on *what I had to know* to arrive at my solution.

My sense of the situation.

I have already spoken of my sense of the situation at a low level of resolution (Fig 1). Now, I had to picture the situation at a higher level of resolution. This was easy, and the result is illustrated in Fig 2.

Now, Fig 2 may seem too obvious to mention. In fact, it is tempting to brush it aside saying, "Every knows *that* !" Remember, however, that without such a body of "common knowledge" the book problem would be much more difficult. In fact, if I was initially unable to frame the mail-order transactions, I would have to learn to do so before I could begin to solve the book problem.

My sense of possible questions to pursue.

Once the situation had been framed along the lines of Fig 2, it was easy to think of several questions that might be pursued (Table 1). When we are familiar with a situation, it is easy to generate lists of questions like this one. But why is this the case?

Table 1. A partial list of questions that might have been pursued at the outset of the book problem.

• Was the money order redeemed?
• Was the money order lost in the mail?
• Was the book lost in the mail?
• Did the bookseller receive the money order?
• Did the bookseller ship the book?
• Who is the bookseller?
• Does Shipping and Receiving have the book?
• Did I actually mail the money order?
• Do banks handle money order transactions like the handle personal cheques?
• What is the telephone number of Shipping and Receiving Department?

I think it has to do with the way we use "frames" in our conceptual thinking (Hofstadter 1979, Minsky 1985). For instance, if I say "professional football player," you are able to form a mental picture of a professional football player. One explanation of this phenomenon goes like this: Your mind contains a frame for "professional football player." This frame is embellished with slots to hold various attributes of the football player: name, height, weight, sex, race, age, position, team, ..., and so on.

If a specific, known, football player is indicated, say, O. J. Simpson, then the slots are filled with that player's known attributes. If, on the other hand, only the generic

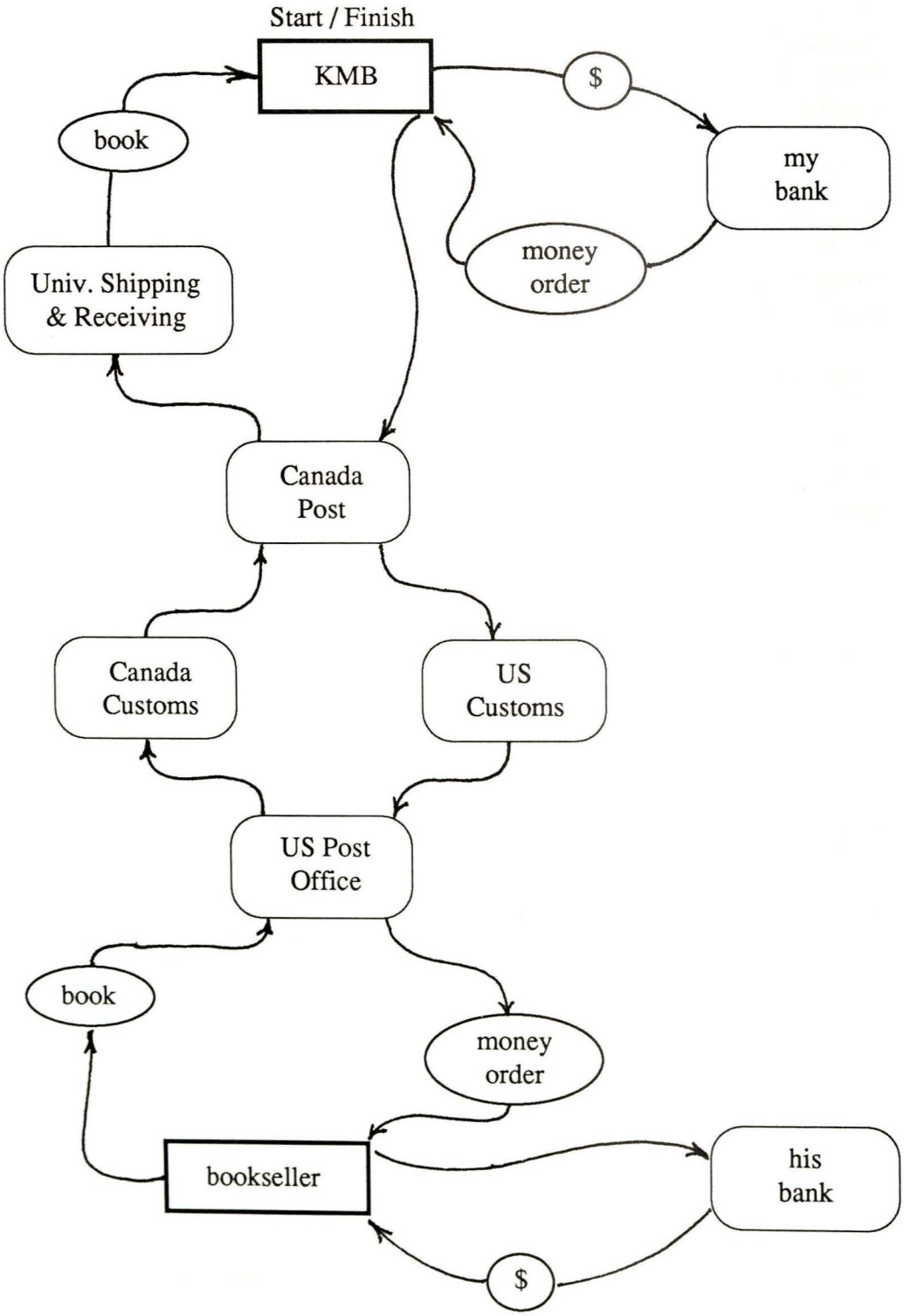


Fig 2. My second model of the mail-order book transaction.

label is given, then some of the slots may be empty (e.g., the name and team slots), and other slots will contain default values.

The default values allow us to form a generic mental picture even when all we have to go on is the generic label "professional football player." My generic football-player-frame has roughly these default values:

name:	_____
height:	tall, say, 6 feet 5 inches
weight:	heavy, say, 250 pounds
sex:	male
race/national origin:	white or black North American
age:	about 30 years old
position:	_____
team:	_____

Of course, *your* generic football-player-frame may be quite different from mine. In fact, your frame may correspond to what I call a "soccer player."

The heads of insiders are full of *domain*-approprate frames, and these frames are embellished with all sorts of *frame*-appropriate slots.

Now, here is how I think our attention is drawn to questions during problem analysis. When we try to use a frame to represent a problematic situation, and some of the critical slots are empty, or hold only default values, we recognize this fact and the need to fill these particular slots. Questions that are useful for defining situations will be of the who-, what-, when-, where-, why-, and how-type (Wales et al. 1987).

My initial frame for the book transaction was quite vague, and it contained many empty and default-valued slots. For instance, under "bookseller" the slots for business name, surname, address, and telephone number were all empty. Since these data were potentially useful, I was prompted to ask: Who is the bookseller? What is his address? What is his telephone number?

My sense of which questions to pursue.

If I had studied the book transaction closely enough, I might have discovered thousands of missing details. For any particular problem, however, only some of these details are critical. For instance, I don't know the name of the Canadian postman who first handled the envelope that contained the money order. I *might* have set out to discover the postman's name, but in fact I did not even consider doing so. That information was never called for during the solution of the book problem. So how do we target questions that are worth asking?

It seems to work like this. First we try to establish a rough sense of the overall situation, a situation-frame (e.g., Fig 2). Then we scan that broad scene and ask, "How might I get from what is to what I want?" I will have more to say about this operation in a moment. For now, suppose that several alternative lines of inquiry come to mind. Then, I think we evaluate each alternative line of inquiry along 2 dimensions. In effect, we ask:

1/ How much will an answer to this line of inquiry contribute to the solution of the whole problem. That is, how *worthy* is this question?

2/ What is the probability that this line of inquiry will be successful? That is, how *feasible* is this question to answer?

The utility of any particular line of inquiry is the product of its worthiness times its feasibility

utility of question = value of answer x probability of success

In common sense thinking, we don't actually do any arithmetic. We rely instead on educated guesses of the utility of various courses of action. Nonetheless, the logical basis for such reckoning is quanitative.

Now we can see why asking about the postman's name would seem *silly* even if that question *had* come to mind.

1/ the value of the knowing his name seems very low
2/ the probability of finding his name seems moderate (?)

On the other hand, asking Shipping and Receiving whether they had seen the book *did* come to mind (because of their strategic position in my situation- frame), and pursuing the question seemed *wise* because

1/ the probability of finding the answer seemed high
2/ the value of the knowing the answer seemed high

Now, let's go back and pick up on a point I mentioned a moment ago. Once we settle on a broad course of action, we try to discover a chain of specific actions that connect what we want to what we have (some thing or bit of information that is within our grasp). This action-chaining operation is much like finding a series of stepping stones that form a route across a stream. We can chain in either direction: forwards or backwards. Here is a simple example of backward chaining:

What do I want? An answer to the question, "Has my book arrived on campus?"
Who might know the answer? Shipping and Receiving
How do I communicate with them? Call them on the telephone.
How can I discover their extension? Look in the directory

If at this point I have immediate access to a directory and a telephone, I have reached the point where I can act. Of course, reasoning at this level is so automatic that it may be completely transparent to the problem solver.

Problems within problems - another kind of structure.

In complex problems, when we attempt to build action chains at the surface level we often discover new problems within the chains. When this happens, problem analysis has to be applied recursively - first at the surface level, and again at each deeper level where "interior" problems exist. Regardless of the level in the hierarchy, the analytical process is the same

1/ define the situation
2/ set a goal based on an assessment of the relative utility of alternative courses of action
3/ find a chain of actions to achieve that goal

In the book problem, an interior problem surfaced as soon as my attention turned to the bookseller. Since the frame slots for the bookseller's name, address, and telephone number were blank, I had to seek ways to fill these slots. When I tried to do so, I found that I was temporarily blocked. I was fortunate, however, that my knowledge of all sorts of things was sufficiently complete and well structured to allow me to succeed. The main features of the find-the-bookseller problem are shown in Fig 3.

The importance of personal experience.

Much of what I needed to know to solve the book problem is part of the common knowledge of every adult North American. For example, we know how to use the telephone directory and directory assistance when a telephone number is sought. The idea to call the clerk in the Northford Post Office, however, might not have occurred to everyone. Why did this occur to me?

I am sure I thought to call the Post Office because I grew up in a small town in Michigan. One of my few links to the bookseller was his P.O. Box, and it occurred to me that the postal clerk might know Mr. "Elliot." Furthermore, I felt reasonably sure that a postal clerk in a small town like Northford, Connecticut would be friendly and helpful - they are in Lake Orion, Michigan. So, here is another example of how useful it can be to have an insider's point-of-view. In this case, I was an insider because of personal experience in an analogous situation.

THE IDENTIFICATION AND ANALYSIS OF RESEARCH PROBLEMS

Problem solving, including scientific research, is hard under either of two circumstances:

- the problem solver is an outsider with respect to the problematic situation
- the problem solver's general problem solving skills are weak

When beginning graduate students suffer from either of these difficulties, the thesis research is not likely to go well. The question is, what can faculty mentors do to help students overcome these problems?

Helping Students Become Insiders

If the student is effectively an outsider with respect to the problem situation, both student and mentor should assign a high priority to overcoming this deficiency. This may be a difficult and time consuming task. The student must assimilate a deep and connected body of knowledge about both the study system and the special research methods of the research discipline. To acquire this knowledge there is no substitute for first-hand experience. Of course, other souces of information are also important, e.g., lectures, reading, and consultation with other researchers, but first-hand experience is essential.

Analogy is a short-cut to becoming an insider. Help students to understand this, and help them find situations in their own experience that can be exploited in the present situation.

Perhaps the most powerful frame for organizing a complex body of infomation is the general concept of system. By the system-frame, I refer to the concept that all sorts of situations can be thought to consist of

- structural components **PLUS**
- an environment **PLUS**
- a web of structural relationships of 2 broad kinds
  component-to-component relations and
  component-to-environment relations

If the student has not completely assimilated the system-frame then introduce it, and frequently demonstrate how you use it to frame problematic situations. [See Churchman (1968) and Laszlo (1972) for a balanced, non-mathematical introduction to the systems concept.]

Two other powerful ideas are concept mapping and the knowledge-V heuristic. These are the inventions of Novak and Gowin (1984).

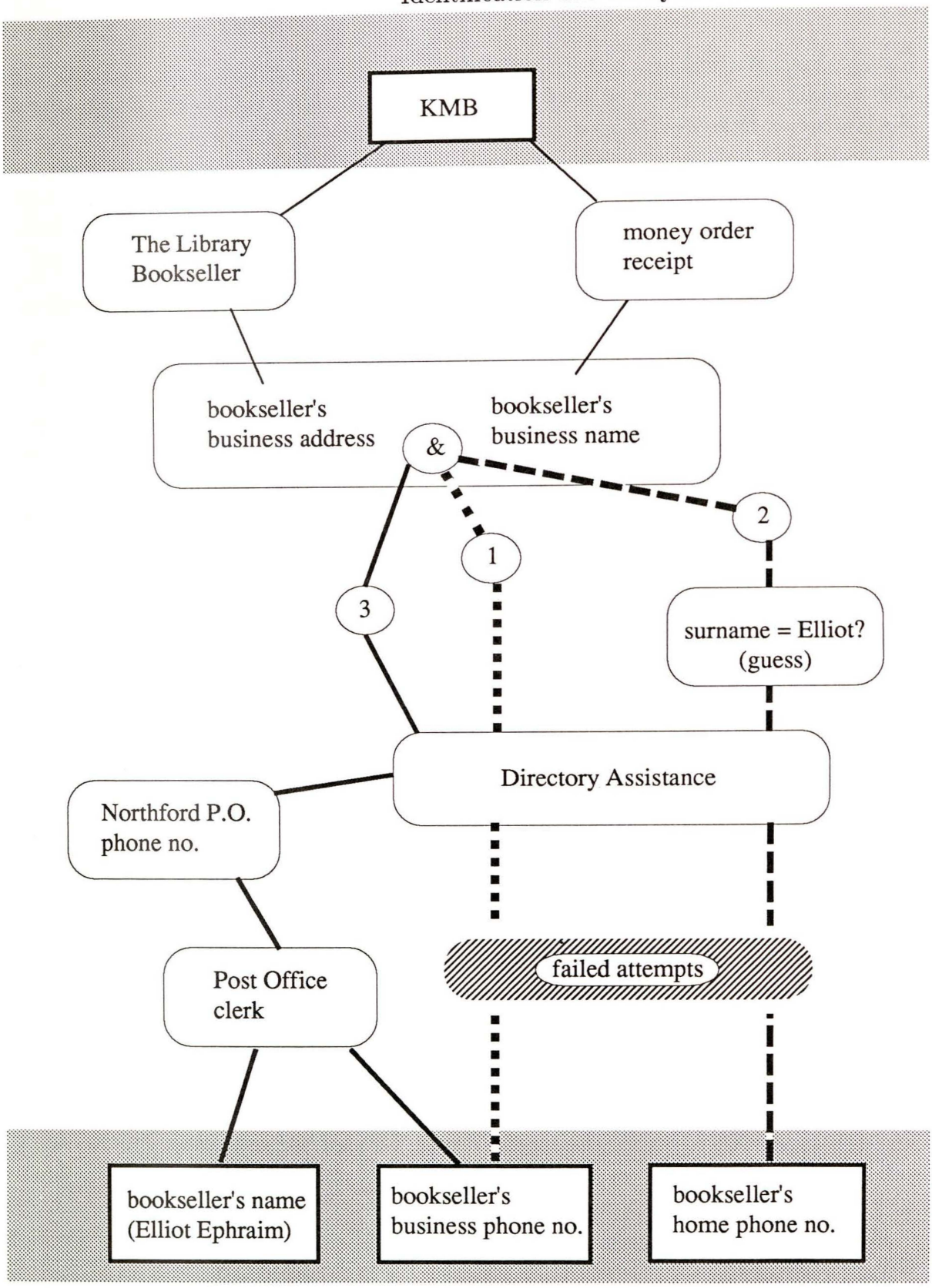


Fig 3. Some of the"stepping stones" used to discover Mr. Ephraim's name and business telephone number. Attempts 1 and 2 were unsuccessful.

Finally, both student and mentor must remember that mastering any difficult field may require several years of hard work and appropriate experience. Don't expect the transition from novice to expert to be either quick or easy.

Helping Students Learn General Problem Solving Skills

Many books have been written about teaching problem solving (e.g., Brown and Walter 1983; Polya 1957, 1962; Resnick 1987; Schon 1987; Whimbey and Lochhead 1982; Wales et al. 1987). I will note just a few highlights here.

Thinking is a skill, like piano playing or swimming. It is best taught, as other skills are taught, by having the student practice under the guidance of a coach. The trouble is, thinking is an invisible skill. Both student and coach must must recognize this, and they must work together to surface and verbalize their thoughts (Whimbey and Lochhead 1982, Schon 1983, Wales et al. 1987). In other words, both must think out loud, and both must be free to ask questions when verbal clues are not forthcoming.

Here are a few specific ideas for faculty mentors to consider when working with new graduate students.

• Make sure the student is able to recognize a good research problem. *Good* problems are worth solving, they are feasible, and, from the researcher's point-of-view, they are interesting.

• Choose a general problem solving strategy, explain it to the student, and refer to it frequently and openly during problem solving sessions with the student.

My choice is Wales' et al.'s (1987) framework for open-ended problems, although there are many others. Wales and his colleagues analyze problem solving into the following stages: define the situation, set a goal, generate ideas, make a plan, and take action. Their first 3 stages are what I have been refering to as "problem analysis."

• Model the way you think during problem solving. This means that the student has to be present when you (the mentor/coach) are solving problems. Furthermore, you must be open about your problem solving activities. You must think out loud.

It is especially important for the student to see how you frame problem situations and how you reason during problem solving. Novices do not realize how much time experts spend reflecting on process. Therefore, you have to be completely open about what is going on. [To learn more about how experts reflect in action, see Boothroyd 1978 and Schon 1983.]

• Set practice problems for the student. If these also contribute to the process of becoming an insider, so much the better.

Good problems are *genuine* problems (from the student's point-of-view) as opposed to exercises, and they are within the grasp of the student. If the solution is out of reach, the problem will discourage rather than instruct. Try to structure the problems in a series so that experience gained with early problems is built upon in later problems.

• Either
  - be there when the student is engaged in problem solving, or
  - discuss the problem with the student after some progress has been made

During these sessions, encourage the student to reflect on the problem solving process. What assumptions were made? How was the situation framed? What goals were set? Why did these goals seem worthwhile?

LITERATURE CITED

Boothroyd, H. 1978. Articulate intervention. Taylor and Francis Ltd, London. 154 p.

Brown, S.I. and M.I. Walter. 1983. The art of problem posing. Lawrence Erlbaum Associates, Publishers, Hillsdale, NJ. 147 p.

Churchman, C.W. 1968. The systems approach. Delacorte Press, New York. 243 p.

Hofstadter, D.R. 1979. Godel, Escher, and Bach. Basic Books, New York. 777 p.

Laszlo, E. 1972. The systems view of the world. George Braziller, New York. 131 p.

Minsky, M. 1985. The society of mind. Simon and Schuster, New York. 339 p.

Novak, J.D. and D.B. Gowin. 1984. Learning how to learn. Cambridge University Press, New York. 199 p.

Polya, G. 1957. How to solve it, 2nd ed. Princeton University Press, Princeton, NJ. 253 p.

Polya, G. 1962. Mathematical discovery: on understanding, learning, and teaching problem solving, vol's I and II. John Wiley and Sons, New York.

Resnick, L.B. 1987. Education and learning to think. National Academy Press, Washington, D.C. 62 p.

Schon, D.A. 1983. The reflective practitioner. Basic Books, New York. 374 p.

Schon, D.A. 1987. Educating the reflective practitioner. Jossey-Bass, Publishers, San Francisco. 355 p.

Wales, C.E., A.H. Nardi, and R.A. Stager. 1987. Thinking skills: making a choice. Center for Guided Design, West Virginia University, Morgantown, WV. 162 p.

Whimbey, A., and J. Lochhead. 1982. Problem solving and comprehension, 3rd ed. The Franklin Institute, Philadelphia. 343 p.

QUALITY OF MODELLING:

UNITING OPPOSITES

Fedor N. Semevsky
The Natural Environment and Climate Monitoring Laboratory
of the USSR State Committee for Hydrometeorology and the USSR
Academy of Sciences, Moscow, USSR

Boris Zeide
Department of Forest Resources
University of Arkansas Agricultural Experiment Station
Monticello, Arkansas 71655
USA

SUMMARY

Trade-offs between model characteristics (accuracy, generality, robustness, and so on) should not be accepted as inevitable. Good models manage to combine desirable features, and such a combination constitutes the quality of a model. Models are not necessarily inferior to reality. After all, it is the model and not the raw data that makes sense out of the world around us. A good model should reflect the overwhelming importance of reproductive effort to living beings. The quality of a model can also be judged by the constancy of its parameters and the absence of patterns in residuals. Conventional statistics characterize quality when the compared models have the same number of parameters or when models are used for extrapolation rather than for interpolation. Straddling a problem by combining two opposite hypotheses for the same process is one of the ways to construct a good model.

Keywords: criteria of model quality, growth models, modelling strategies, reproductive effort

A NEGLECTED FEATURE OF MODELS

Models are characterized by accuracy, generality, complexity, testability, robustness, flexibility, and many other attributes. It is commonly believed that "good" characteristics cannot be combined in one model: accurate models lack generality, simple models are not realistic and so on. The view that such trade-offs are inevitable was eloquently expressed by Levins (1966) who outlined several strategies for model building in population biology. These strategies are distinguished by "sacrifices" among generality, realism, and precision. The view that these sacrifices

cannot be avoided is accepted in ecological (Kareiva 1989) and forestry (Sharpe 1990) modelling.

Levins' strategies imply and even sanction a greater sacrifice, the advancement of science. The trade-off between model characteristics is a hallmark of mediocre, expendable models. These "strategically" built sacrificial models are irrelevant to scientific progress. The most pertinent feature of a model is its quality. Although this feature is missing from even the longest lists of model characteristics, it should never be sacrificed.

A good model does not trade accuracy for generality or vice versa but encompasses both. Understanding is advanced by models that are accurate and at the same time are simple and general. If a model can be depicted as a point in space with axes representing increasing accuracy, generality, robustness, etc., then model quality can be quantified as the distance between this point and the origin.

Plants and animals often resort to trade-offs and sacrifices. These activities alleviate deficiencies of light, water, nutrients, and other factors. They promote fitness and contribute to the survival of the organisms. These trade-offs should not be confused with those in models in which trade-offs reflect the inability to grasp reality. It is not necessary to describe deficiencies of nutrients with deficient models.

Contrary to Levins' (1966, p. 422) belief that "this cannot be done," models that excel in many of the above aspects do exist. Levins' own contributions, aside from his philosophical digressions, may testify to this. Although rare at any given period, these superior models constitute the majority, if not all, of models that have shaped our knowledge. In the process of natural selection of models, only those of exceptional quality find a place in the genealogy of science. The model of the solar system by Copernicus and Kepler, to use a familiar example, is more accurate, simple, and meaningful than that of Ptolemy. Sacrifice and trade-offs can and should be avoided in modelling.

MODELS AND REALITY

Models are sometimes viewed as necessarily inferior to reality. Again, Levins (1966, p. 430) expressed this view best: "All models leave out a lot and are in that sense false, incomplete, inadequate." As a model of models, this sentence is self-defeating. Besides this logical problem, Levins' view is inadequate because of the intimate connection between a model and the process of data acquisition. Without a model, we simply cannot observe or collect data. If models were indeed false, so would be the data. We would be caught in a vicious circle between false models and misleading data.

To show that models can be superior to reality consider, for instance, the process of uninhibited cell division, represented by the sequence 2, 4, 8, 16,.... It could be modeled by a high degree polynomial. A polynomial with the number of parameters equal to the number of data points would match the sequence exactly, even when the numbers deviate from doubling. Given these deviations, the exponential function, the only sensible model of the process, will be less accurate. Still, the exponential, and not the polynomial, is preferable. The proper function, in this case the exponential, screens random noise and presents nature in its refined form. Precisely because this function, in Levins' words, "leaves out a lot" (in this instance, noise), it is correct and adequate, while the indiscriminating polynomial that exactly copies the raw data is irrelevant.

The work of a scientist is a constant struggle to filter out chance deviations and penetrate beyond appearance. The Latin source of our word "understand" (*intelligere*) means to read what is inside a thing (*intus legere*). With all due respect for facts, there are many situations when, given contradiction between facts and a model, one would say so much the worse for the facts. Although it might sound contradictory, this is an expression of the highest regard for facts and not contempt of them. At every turn of model construction, we are dealing with two uneven groups of facts: those few at hand and the total sum of knowledge accumulated by humans. If a model is derived from basic axioms that embody this sum of facts, then we might trust the model rather than particular facts.

Another dichotomy of modelling deals with two different faces of reality, the past and the future. Growth models aim at predicting growth 5, 10, 20 and more years into the future. Certainly, actual future growth will be a more accurate mirror of itself than that produced by a model. But the future is inaccessible.

The existing data describing growth of trees in the past represent another part of reality. Can we use data from one stand to predict future growth of another stand? We can, but not directly. Usually, we analyze the data to determine to what degree the soil, location, stand history and other factors are similar to those of the investigated stand. To facilitate the analysis and reduce the information to a manageable number of variables, we fit a model to the data. Even though there is some justification for viewing models as false and inadequate, in our work we prefer them to the reality, which is either inaccessible or only remotely relevant.

To sum up, quality is the chief characteristic of models. It should be neither traded away nor sacrificed. Good models manage to combine desirable features and, in some respects, are preferable to reality. These models, not raw data, make sense out of the world around us. Good models describe and predict reality better than those "strategically" built models. Yet, the main difference among models is that quality models contribute to understanding of nature while others do not.

CRITERIA OF QUALITY

Superior models, such as Mendel's genetic combinatorics, attract fresh minds by their logic, coherence, and beauty (in short by their quality) even when empirical evidence is weak or missing. A question arises as to whether there are more tangible criteria of quality besides this nebulous attraction to fresh minds and beauty.

The fact that an irrelevant formula (for example, a polynomial) can describe a process more accurately (that is, with a smaller standard error or sum of squared residuals) than the proper function means that routine statistics are not always are reliable criterion of quality. Several more appropriate criteria are discussed below. These criteria of quality can be divided into two groups. One group, illustrated by the first of the following criteria (Maximization of reproductive effort), addresses coherence or reasonableness of models. The second group deals with the correspondence between model and reality.

Maximization of Reproductive Effort

At first glance, as a general criterion of ecological modelling, the maximization of reproductive effort looks odd. In trees, seed production is trifling, periodicity is irregular, and measurements are not reliable. Currently, only a few models reflect reproduction in an auxiliary submodel. There is a general, if tacit, consensus that the effect of reproduction on tree and forest growth is negligible and can be safely disregarded. Thus, the index of the most recent and large collection of works on forest growth modelling (with 38 contributions) contains no entry on reproductive effort or seed production (Dixon et al. 1990). For humans interested in the production of wood, oxygen, or clean water, this consensus is expected.

For trees, however, the items we value (foliage, branches, stems) are merely means of maximizing reproduction. Models neglecting the overwhelming significance of reproduction reflect our vested interests rather than those of the tree. In a model, plant growth and all other activities should be presented as a way to maximize reproductive effort. A model attempting to describe and explain biological phenomena from the tree's standpoint is likely to be more meaningful and successful than one based on our values. Instead of a mere description of growth or mortality, any forest process or pattern should be viewed as a contribution to the ultimate goal of tree life.

A focus on maximizing reproductive effort would not necessarily make models more complicated. To estimate reproductive effort, it is not necessary to count every seed. Reproductive effort can be evaluated with information already at hand. This effort is equivalent to the difference between the efforts expended for growth and the returns in the form of assimilates. The parameters of a model should be calculated

so as to maximize this difference. Such a model would be more efficient because the condition of reproductive maximum is equivalent to knowing one parameter of the model.

This approach sometimes leads to novel and unexpected results. For example, plant growth is usually pictured as a smooth curve that can be differentiated at any point. The smoothness of this curve is a secondary phenomenon that masks a break along the curve. This break indicates the onset of reproduction. As was proven mathematically, in a predictable environment maximal reproductive effort occurs only with the complete switchover from vegetative growth to seed production (Cohen 1971, Insarov 1975). The lack of environmental predictability smooths the transition. Actual growth curves result from a compromise between maximization and acceleration of reproductive effort.

When growth is considered as a way of maximizing reproductive effort, growth curves can provide much more information than is realized. They can reveal the intensity of competition responsible for timing of the transition, the expected longevity of plants, and the degree of environmental predictability. Assuming that this degree is constant on a given site, the smoothness of a growth curve can tell us about the ability of various plant species with similar phenology to anticipate changes in the environment.

The application of the described criterion is simple. When considering a model, we should check whether it is designed to maximize reproductive effort. A fuller presentation of this approach to modelling along with many applications is given in Semevsky and Semenov (1982, reviewed in Zeide 1986). This approach is based on the optimality principle, which is a generalization of the theory of natural selection.

Constancy of Parameters

A model is supposed to explain a process and expose its essential features. One of these features is the change of the dependent variable with respect to independent variables. The predicted values must change to match the modeled variable.

Another complementary feature uncovered by a model is the constancy of parameters that govern the change of the variables. We study not the change but only its repetitive patterns. A change without a pattern is beyond science. A model presents the means to discover these patterns. By tracing the apparent change, a model reveals hidden invariance. Therefore, parameter constancy might be a more reliable measure of quality than minimum of squared errors.

This constancy can be tested by comparing values of the parameters calculated at different segments of the model's domain. The above example of cell division can

be used to demonstrate this procedure. Given two segments of, for instance, five points, the exponent will be less accurate on each of the segments than a fifth-degree polynomial. However, unlike the polynomial, the parameters of the exponent will remain practically the same on both segments.

Lack of Patterns in the Residuals

Reality, as presented by raw data, contains two opposite elements: meaningful pattern and misleading noise. The quality of a model is determined by its ability to separate these elements of reality. Technically, this aspect of quality can be judged by the absence of patterns in the residuals. This is possible when the model fits exactly the underlying pattern (not the data). It seems that the total amount of order is invariant. When order is present in the residuals, it is lacking in the model and vice versa.

Least Squares (With Qualifications)

Under certain conditions, conventional statistics based on least squares can be of help when judging a model's quality. When compared models have the same number of parameters, the more accurate one would be of higher quality. Least squares also are appropriate for describing quality of models when they are used for extrapolation rather than for interpolation. For example, in the case of cell division, an exponent will predict the future number of dividing cells more accurately (in terms of least squares) than a polynomial of any degree.

FITTING A MODEL

The application of least squares for the estimation of parameters and evaluation of the accuracy of models is justified when the assumptions of this method are satisfied. Among these assumptions are: (1) the exact (functional) model of a process is known; (2) the domain of the parameters is unknown. For example, the average height of a stand can be anywhere between $+\infty$ and $-\infty$ with equal probability; (3) minimization of the sum of squared deviations from the predicted values is the best way to approach true values.

Too often in forest practice none of these assumptions are warranted. As a rule, we do not know the functional model. On the other hand, we do know that the parameters of interest are located in a finite range. Thus, foresters can easily estimate average tree height by eye within a pretty narrow range. In fitting a model, we are interested in minimizing the distance from observed to true values rather than to predicted values such as the sample mean.

A priori knowledge of the expected values of parameters and their range makes it possible to design more efficient estimates than those produced by the least squares method. A proof that such a possibility exists is provided here for the simplest case, which deals with the estimation of one parameter, the population mean.

According to the least squares method, the arithmetic mean of a sample of n objects, $\overline{x} = (\sum x_j)/n$, is the best estimate of the true (population) mean. x_j might denote a particular measurement of tree j, such as its height or basal area. A more efficient estimate of the true mean, μ, of the entire population uses the weight k, which is different from unity and specific for a given population. The proof is based on a new criterion of minimization. Instead of minimizing the sum of squared deviations from the sample mean, we minimize the deviations from the unknown true mean, μ:

$$\sum(kx_j - \mu)^2 \rightarrow minimum \tag{1}$$

The factor k, found from this condition:

$$\frac{d[\sum(kx_j - \mu)^2]}{dk} = 0 \tag{2}$$

is equal to:

$$k = \frac{\mu \sum x_j}{\sum x_j^2} \approx \frac{n\overline{x}^2}{\sum x_j^2} < 0 \tag{3}$$

In general, to estimate parameters of any function so as to minimize the distance between the unknown true values and their approximations, it is necessary to minimize the following functional:

$$F = \int_0^\infty y_m(\chi)d\chi \int_A q_s(Y_i - f_\alpha(x_i^*))p(\alpha)d\alpha \int_\Omega ABS(f_\alpha(x) - f_a(x))dx \tag{4}$$

where $y_m(\chi)$ - distribution of sample variance with m degrees of freedom;
x - vector of independent variables;
x_i^* - vector of observed independent variables;
$f_a(x)$ - an approximating model;
$f_\alpha(x)$ - actual (true) values of the function;
Y_i - observed values of the function;
a - vector of parameters' estimates;
α - vector of possible values of parameters;
$p(\alpha)$ - a priori probabilities of the parameters' values;
$q_s(Y_i - f_\alpha(X_i^*))$ - observed probabilities of the parameters' values;
A *and* Ω - range of integration determined from information on the parameters' range;
ABS - absolute value.

By utilizing known information about the studied process, this functional makes it possible to estimate its parameters more efficiently.

HOW TO CONSTRUCT A GOOD MODEL

Although there is no rational answer to this question, experience gained through oscillations from one model (or "paradigm") to another can provide certain hints for model construction.

The separation of reality into a comprehensible pattern and irrational noise might be misleading because we do not know where the pattern is hidden and whether it is unique. Too often we tend to assume uniqueness. Given a relationship between variables, which might be visualized as a scatter of dots, we immediately and almost instinctively tend to draw a single line representing the central trend. This line is presumed to convey the essential information of the data, free from accidental variation. The same is true in more complicated situations that cannot be reduced to a cloud of dots. We tend to concentrate on the unique essence represented by the mean and believe in a single explanation of a given event at the expense of the entire range of possibilities.

It is useful to define the same approach in negative terms: we are reluctant to admit the coexistence of opposing explanations and allow contradictions within a single entity. Facing a diversity in reality, we respond by separating the opposites either in space (by dichotomizing objects or producing conflicting contemporaneous theories) or in time (by holding contradictory theories consecutively), by seeing black or white but not both. We try to keep a given event uniform and free from contradictions, either ordered or disordered, animate or inanimate, deterministic or stochastic, constant or variable, yes or no, good or evil, and so on. An appeal to formulate alternative hypotheses (Platt 1964) does not contradict this approach because the intention is to weed out all but the best one.

Experience, both within and outside biology, indicates that no single proposition is fully satisfactory and that our thoughts tend to oscillate between opposites. In our attempts to pinpoint a problem, we too often miss the point. A more reliable approach would be to straddle the problem before pinpointing it. If indeed we can learn from our experience, it makes sense to cut through the zigzags and first formulate simultaneously two idealized extreme hypotheses of the same process and only then proceed to model the central tendency. In terms of a dot scatter this means that two lines representing extreme cases are drawn before the medial one. The two extreme hypotheses or models might separately describe the completely random and totally ordered cases, density-dependent as opposed to density-independent processes, predictable versus unpredictable environments, r- and K-selection, territorial or group behavior, biotic potential and environmental resistance, and the like.

Two extremes are clearly seen in the ways plants maximize their reproductive ef-

fort. Natural selection favors those plants that reach the optimal compromise of two opposing strategies. The first is to maximize reproduction by delaying it. This has proven to be the best strategy in a predictable environment. In a totally unpredictable environment, the opposite strategy would be of selective advantage: accelerate reproduction, at the expense of seed quantity, to assure some seed in case of sudden death. Actual growth results from some combination of both strategies. Although these strategies are analytical products of our mind, their action has been demonstrated experimentally (Zeide 1978, Wolgast and Zeide 1983).

Our experience in modelling individual survival of trees provides another illustration of the described approach. At one extreme, it can be assumed that the size of trees in a given stand varies at random and is independent of tree position and past interaction with neighbors. Based on this assumption it was proven that the survival probability of a tree, p_i, is a power function of its normalized size, or rank, r_i (Zeide and Semevsky 1972):

$$p_i = r_i^v \tag{5}$$

The exponent v is the average number of "victories" won by each surviving tree. It can be calculated as the ratio of the number of trees that died to the number of those that survived during a certain period:

$$v = (N_0 - N_1)/N_1 \tag{6}$$

where N_0 and N_1 are the numbers of trees at the beginning and the end of the period.

At the other extreme we can assume that the variation in tree size results exclusively from the interaction between trees, so that only the largest trees will survive:

$$p_i = \begin{cases} 0 & \text{when } 0 < r_i \leq m \\ 1 & \text{when } m < r_i \leq 1 \end{cases} \tag{7}$$

where $m = (N_0 - N_1)/N_0$ indicates the proportion and the highest rank of trees that died (mortality).

Actual size of trees results from both past suppression and random events. Therefore, on a plane with the axes representing tree rank and survival probability, the observed survival of trees lies in the field bounded by the lines described by equations (8) and (10). In addition to the perimeter of possible solutions, these equations provide the following boundary conditions that specify the position of the line depicting the expected actual survival. Because the extreme lines intersect at the point ($r_i = m, p_i = m^v$), the expected line will pass through this point. The second derivative of this line will be zero at the same point, while the first derivative will be zero when $r_i = 1$. The integral of the line must be equal to N_1/N_0. Thus,

the consideration of imaginary ideal cases, which seemed quite remote from reality, proves to be helpful in formulating a realistic model of tree survival.

There are several advantages to modelling based on extremes. Consideration of two coexisting opposites promotes understanding of the resulting phenomenon. It is often much easier to describe two idealized situations than a single actual one. Mutual consideration of two extreme models usually provides substantial information (such as initial values, position of the inflection point, and other boundary conditions) for modelling the central tendency. Variation becomes bounded so that actual events can be viewed as participants in both extremes, instead of just being random deviations. The focus on extremes, rather than the mean, alerts us to the emergence of a new trend that can be visualized as branching from the original.

The major advantage on which the above mentioned points are based is that many biological phenomena actually result from the combined action of two opposing factors or groups of factors. Ultimately, all change in ecosystems results from the conflict between infinity implicit in multiplicative reproduction and the limit imposed by finite space. Therefore, it is only natural that good models reflect this conflict and its resolution.

The outlined approach is not an unerring recipe for model building. It is not easy to uncover simultaneous extremes and incorporate them into a model. Ecological modelling will always be a challenge to our creativity.

ACKNOWLEDGMENTS

We are grateful for valuable comments made by John L. Greene, Daniel J. Leduc, Lynne C. Thompson, and Suzanne Wiley.

LITERATURE CITED

Cohen, D. 1971. Maximizing final yield when growth is limited by time or by limiting resources. Journal of Theoretical Biology 33:299–307.

Dixon R.K., R.S. Meldahl, G.A. Ruark, and W.G. Warren, Editors. 1990. Process modeling of forest growth response to environmental stress. Timber Press, Portland, Oregon. 441 p.

Insarov, G.E. 1975. Stepped model of growth and reproduction of organisms. In Kolichestvennie aspecti rosta organizmov, pp. 114–121. Edited by A.J. Zotin. Nauka, Moscow. 291 p.

Kareiva, P. 1989. Renewing the dialogue between theory and experiments in population ecology. In Perspectives in ecological theory, pp. 68–88. Edited by J. Roughgarden et al. Princeton University Press, Princeton, New Jersey. 394 p.

Levins, R. 1966. The strategy of model building in population biology. American Scientist 54 (4):421–431.

Platt, J.R. 1964. Strong inference. Science 146:347–353.

Semevsky, F.N. and S.M. Semenov. 1982. Mathematical modeling of ecological processes. Gidrometeoizdat, Leningrad. 280 p.

Sharpe, P.J.H. 1990. Forest modeling approaches: compromises between generality and precision. In Process modeling of forest growth responses to environmental stress, pp. 180–190. Edited by R.K. Dixon, R.S. Meldahl, G.A. Ruark, and W.G. Warren. Timber Press, Portland, Oregon. 441 p.

Wolgast, L.J. and B. Zeide. 1983. Reproduction of trees in a variable environment. Botanical Gazette 144:260–262.

Zeide, B. 1978. Reproductive behavior of plants in time. The American Naturalist 112:636–639.

Zeide, B. 1986. The idea of invariance in biology. BioScience 36:494, 496–497.

Zeide, B. and F.N. Semevsky. 1972. A theory predicting the chances of survival of individual trees in a population. In Dokladi. Obschaya biologia. 1970, pp. 105–106. Edited by A.L. Yanshin. Moscovskoe obschestvo ispitateley, prirody, Moscow. 111 p.

Prepared July 5, 1990

COGENCY IN FOREST RESEARCH: II

Rolfe A. Leary
USDA Forest Service
North Central Forest Experiment Station
1992 Folwell Avenue
St. Paul, Minnesota 55108
USA

SUMMARY

As a science matures it moves toward breadth of answers, depth of questions answered, and cogency. Cogency means the ideas, the constructs, of a science are clearly identified and stated, and are organized to the maximum extent possible. Greater cogency is needed if we are to keep abreast of the rapidly increasing forest science literature. Cogency cannot effectively be imposed by another party, such as a user-scientist or knowledge engineer, on another scientist's research, so responsibility should rest with the doing-scientist. Cogency in published literature might be increased by analyzing forestry constructs (concepts, propositions, theories) using methods from semantics and logic. It might be increased in research planning if a completed Gowin's Vee were required for all study plans and summaries. It might be increased in research reporting if scientific journals requested explicit statements of scientific hypotheses in their instructions to authors, much as competitive granting agencies do in their instructions. The lack of cogency in forest research reports may also be related to the personal styles of forest scientists.

Keywords: mature science, research quality, Gowin's Vee

INTRODUCTION

Mario Bunge (1968) says mature sciences have depth, breadth, and cogency. Elsewhere I have argued that sciences answering difficult "why?" questions have greater depth than sciences answering "what is the character of?" questions (Leary 1985a, 1988). Likewise, sciences that produce very general answers, i.e., universal or bounded universal statements, have greater breadth than those that produce answers expressed as singular statements. Where does cogency fit in?

Cogency is defined in Webster's New Collegiate Dictionary (1979) as: "... **2 a:** appealing forcibly to the mind or reason: convincing ... **b:** presented in a way that brings out pertinent and fundamental points ...." Mario Bunge (1968) suggests the following interpretations for "cogency": "Better organized" (pg 121, line 2), "organized axiomatically" (pg 121, line 23), and "logical tightness" (pg 144, line 6).

So, it is the ideas or constructs of a discipline that are better organized in cogent science, not, for example, methods or procedures.

THE NEED FOR COGENCY

The need for cogency in forest research can be better appreciated when we consider:

1. The job of a research organization is to discover, organize, deliver, and evaluate new knowledge and technology about what it is charged with researching (Leary 1989).
2. It is possible, although politically unwise, for a research organization to focus all its efforts on knowledge discovery, while slighting organization, delivery and evaluation of knowledge and technology.
3. The point will come when what is "known" must be organized in some fashion before further work can profitably be undertaken. There has been in forest research, as in nearly all areas of science, a very rapid growth in published research over the last decades. The number of pages published in North American forest research journals has sevenfold (10.8 percent annual rate of compound interest) over the past 19 years (Figure 1). If considered on a worldwide basis, the increase would probably exceed sevenfold.
4. While the number of published pages in journals is rapidly increasing, the entire system for making research public in usable form is being stressed. The journal publishers are being strapped financially to pay for publication without significant page charges, and are finding it difficult to ensure thorough review of manuscripts submitted for publication. The scientific community is being stressed to keep up with the mountain of materials arriving daily in journals, proceedings, books, and other forms of communication. Review papers in scientific journals and Forestry Abstracts can help to assemble and focus what is known about a subject. However, there is a difference between what a review article accomplishes and what I interpret "better organized" to mean in the definition of cogency. The user public is being stressed in two directions. They are being required to: a) sift through the explosion of published materials for usable bits, and b) organize the usable bits into larger wholes.

WHO SHOULD DO IT?

Whose responsibility is it to improve cogency — to better organize forest research — to organize forest science findings axiomatically? Four candidate groups come to mind: 1) user-scientists, 2) application-oriented users, 3) knowledge engineers from the artificial intelligence community, or 4) scientists doing research.

One all-too-common attitude among author scientists is that the scientist actually doing research has little time to spend on such niceties as cogency. Spending time

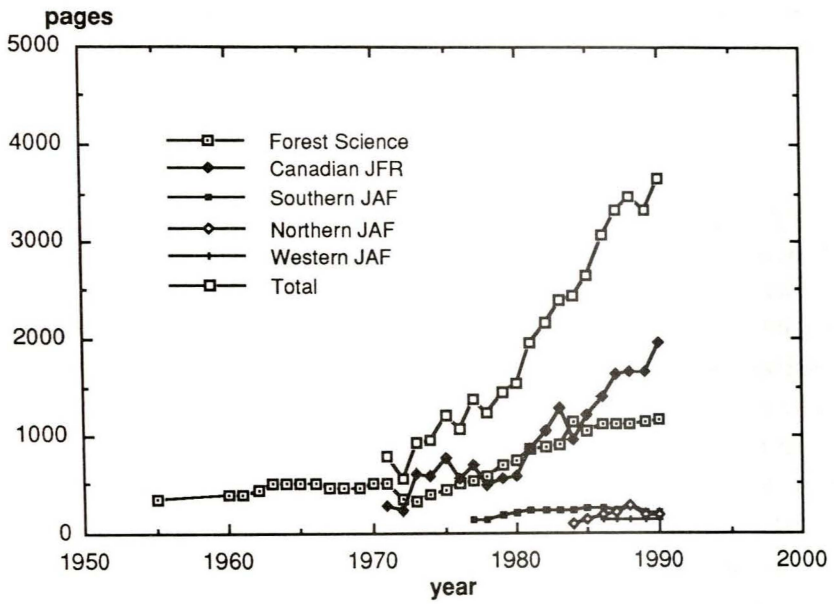


Figure 1. Pages published in North American forest research journals 1955–1989. The Canadian Journal of Forest Research began using a larger page size in 1984. The journals of applied forestry [AJF] (Northern, Southern, and Western) each contain a few pages of advertising which have not been deducted.

on cogency means less time doing research, and actual research time is scarce for many scientists. Furthermore, journals don't have any special requirements about cogency. Efforts at cogency are not common, so some reviewers might even be annoyed by any special efforts in that direction and as a result lower their evaluation of a manuscript. Therefore, there is some incentive for author-scientists to pass the buck to another person.

Research organizations not including organization and delivery components in their mission statements are, de facto, leaving the job to users of their research findings. As a general rule, users are equally short of time, and may be less qualified to anticipate application problems and limitations of findings.

Recent research in artificial intelligence theory and methods, including knowledge organization and delivery systems, along with the development of persons called "knowledge engineers," would suggest that these persons are uniquely qualified for the task of "better organizing" the products of research. Knowledge engineers may have some useful theory and methods to offer scientists; however, it is always dangerous, in my view, for a person other than the doing-scientist to assemble research findings into larger delivery packages. The recent experience of R. Lewis (1980, 1982) in attempting to axiomatize several hundred biological theories suggests it is

discouragingly difficult to isolate the salient ideas that are being tested and reported in published research articles.

Last, we could make the "doing" scientist responsible for cogency in his or her research program and findings. In fact, of course, there is no better person to be held responsible for cogency than the scientist who did the research. Cogency is not just good writing, something that can be incorporated into research reporting — to be significantly helpful in the long run, cogency should be there during study design.

WHERE TO LOOK FOR COGENCY IMPROVING THEORY AND METHODS

Improving cogency requires scientists to look beyond science to semantics, epistemology, and, of course, logic. Contrasting science, semantics, and epistemology is aided by a partial listing of their traits (from Bunge 1974, pgs 196-197):

Trait	Science	Semantics	Epistemology
referents	things (concrete systems)	constructs & signs	knowledge
method	hypothesis & theory, observation and experiment	postulation & proof, and checking with logic, mathematics, and science	analysis, postulation, & checking with substantial knowledge & methodology
goals	finding laws, describing explaining & predicting	elucidation & articulation of concepts, of meaning, truth & cognates	elucidation & articulation of all the concepts about factual knowledge & ignorance
roles	basing technology, controlling world views	conceptual hygiene, spotting genuine referents, clarifying sense, and scotching myths in the philosophy of scientists	methodological alertness & open-mindedness

Of particular importance in semantics are the **roles** of conceptual hygiene, and **goals** of elucidation and articulation of constructs. Improving cogency is a combination of science, semantics, and epistemology.

IN WHAT WAYS CAN COGENCY BE INCREASED?

Cogency can be increased in at least three ways: 1) reseachers can do construct analyses as a regular part of planning studies so that when it comes time to report results, the ideas of the science will have been well organized all along, 2) reseachers can alter the required study plan so that it follows Gowin's Vee, and 3) journals can

alter their submission format to require explicit statement of scientific ideas and hypotheses.

Construct Analyses

While cogency applies primarily to argument, hence to theory, it may not be possible to go directly to axiomatization of theories without first considering what theories are made of — propositions — and before that what propositions are made of — concepts. How concepts, propositions, and theories are made cogent differ, so I will treat them separately, beginning with the most elemental unit, concept.

Concept: A scientific concept is a fundamental unit of thought. It is not a complete thought as would be expressed in a whole sentence. Concepts can be analyzed from a number of perspectives. For example, we can ask, "What does a concept do?" Bunge (1974) suggests it maps objects into statements about objects. A scientific concept bridges the gulf between the world of objects and the world of language, and can be symbolized:

$$P: \{X\} \times \{Y\} \longrightarrow a,$$

where P designates a concept, e. g. , "predation,"
{X} and {Y} designate sets of objects,
× designates the Cartesian product of sets,
⟶ denotes mathematical mapping, and
a designates the word 'predation' (Bunge 1974).

We may also ask, "What kinds of concepts are there?" Listed below is a taxonomy of concept kinds (Leary 1985b, pg 13-14, after Bunge 1967):

Concept kind	Examples
individual	"General Sherman sequoia"; "x"
class	"hardwood"; "softwood"
relation	"is taller than"; "outgrows"
quantitative	"basal area"; "total height"

A definite individual concept is usually designated by the individual's name, whereas an indefinite individual concept is designated by a lower case letter from the end of the alphabet. Class concepts apply to groups of individuals, where membership in the class depends on objects having required properties. Relation concepts deal with relations among objects of some kind, and quantitative concepts function to map elements in the concept domain into the range (typically the set of real numbers).

The logical structure of a concept makes use of what is known about its predicate structure (Bunge 1974). Concept degree refers to the number of objects or sets of

objects the concept maps from the world of objects to the world of language. For example, "hardwood" is a class concept of degree one: Hx or H(x) symbolizes that object x has the qualities needed for membership in class H. Likewise, predicates of degree two relate two individual objects and are symbolized xHy, where H is a degree-two concept. Concept order refers to the nature of the sets of objects related by the concept. For example, the degree-two concept "predation" relates wolf and moose, and can be symbolized wPm, where "w" designates wolf, "m" designates moose, and "P" designates predation. Because the objects wolf and moose are things, this use of the concept "predation" is second degree because two objects are involved, and first order because the relation is between real objects. If, however, the concept is the relation concept "interaction," designated I, the "objects" could be "predation" and "competition," as in pIc. Here, "interaction," I, designates a second order relation concept because the two objects related are first order concepts. Concepts of higher order are more abstract.

Finally, concepts may have different metrical characters: qualitative (good site, low site, better site) or quantitative (ordinal scale, site class — I, II, III; or ratio scale — site index 65).

An example of increased abstraction dealing with concepts representing space occupancy by trees is shown in Figure 2. Subjective "measures" of space occupancy, e.g., "stocking," are class or relation concepts. Objective measures are quantitative. Objective concepts used to represent space occupancy may be absolute or relative. The absolute branch contains first degree, first order concepts. The relative branch contains second degree concepts because the concept relates a subject stand with another stand. The "induced change" branch of the relative branch contains first order concepts. Concepts in the "spontaneous change" branch of the relative branch are second order because both arguments involve properties of trees involved in interaction — one being the null interaction or the open-grown condition.

Action-oriented scientists tend to test the variables (variable letters designate concepts) in mathematical equations and make assessments of variables — and designated concepts — by such measures as goodness of fit of the resulting model to observational data and model predictive potential. Conceptual analyses such as done here help clarify the abstractness of concepts and their utility in efficient expression of models to represent complex phenomena. However, scientists should probably not expect to regularly publish such concept analyses.

In sum, scientists should be able to state all the important new, as well as traditional, concepts in their discipline or subdiscipline (Cherrett 1989), and should know kind, predicate degree and order structure, and metrical character of each. Quantitative concepts having higher degree and order predicate structures offer important economies in expressing factual propositions, are found in advanced sciences, and

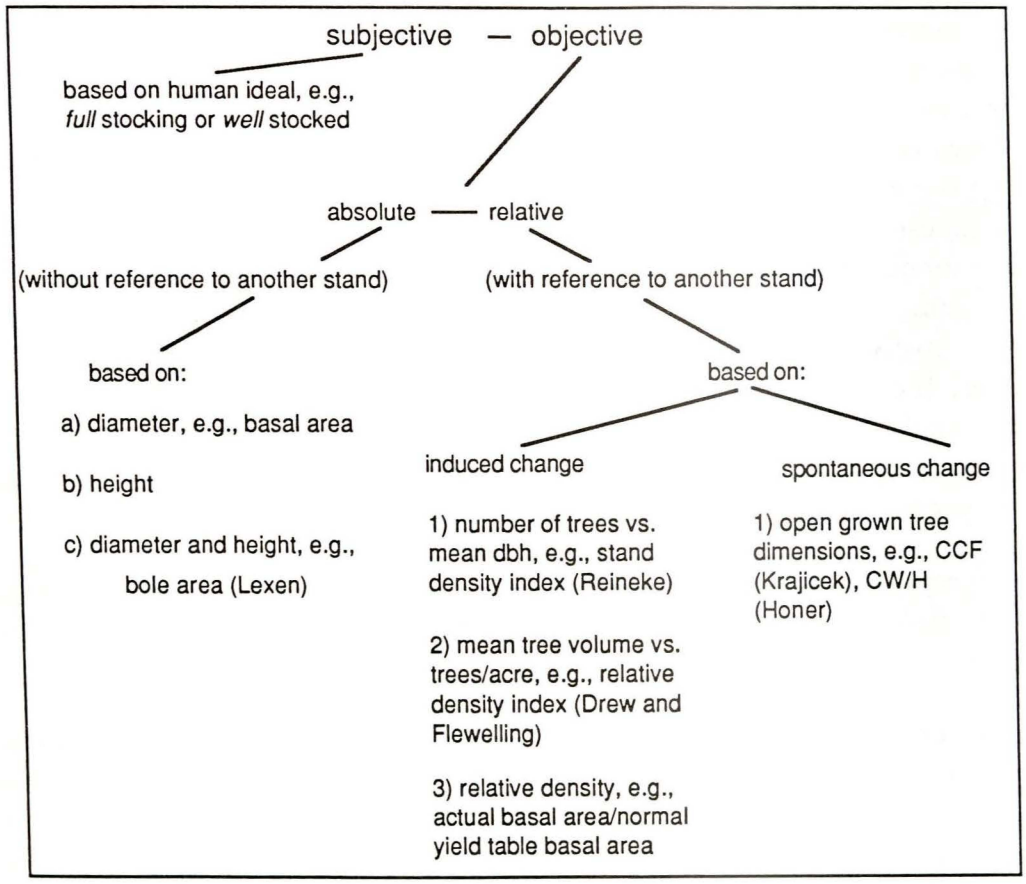

subjective — objective
based on human ideal, e.g., *full* stocking or *well* stocked
absolute — relative
(without reference to another stand)
(with reference to another stand)
based on:
a) diameter, e.g., basal area
b) height
c) diameter and height, e.g., bole area (Lexen)
based on:
induced change
1) number of trees vs. mean dbh, e.g., stand density index (Reineke)
2) mean tree volume vs. trees/acre, e.g., relative density index (Drew and Flewelling)
3) relative density, e.g., actual basal area/normal yield table basal area
spontaneous change
1) open grown tree dimensions, e.g., CCF (Krajicek), CW/H (Honer)

Figure 2: Concepts useful for representing the space occupied by a stand of trees (Leary 1985b). Induced change is Bunge's (1974) term for change resulting from interaction, in this case tree-tree. Spontaneous change is his term for change in the absence of the tree-tree interaction.

should be developed whenever possible in less advanced sciences like forestry.

Proposition: Philosophers agree a proposition is not just a declarative sentence. It is the sentence and its meaning. Propositions may be analyzed following principles laid out for concepts. However, propositions have an additional trait not held by concepts — they may be true or false. Concepts are neither true nor false — they are just useful or not useful. Determining the truth or falsity of a proposition is not germane to cogency analysis, so I will not discuss it here.

A proposition consists of factual and formal concepts. Formal concepts are notions such as "greater than," "belongs to the set," and so on. They are not about a property of a real object.

Propositions may be internally evaluated by looking at:

a) the number of factual concepts they contain,
b) the predicate degree of each concept,
c) the predicate order of each concept, and
d) the metrical character of each concept.

Propositions may also be evaluated externally according to their range. Range refers to what-when-where the statement applies. Ranges drawn from Bunge, using forest research examples, are as follows:

1. Singular: "That tree is a hardwood."
2. Indefinite existential (neither time nor place specified): "There are lots of hardwoods."
3. Localizing existential (time or place or both specified): "There are lots of hardwoods in Buchanan County."
4. Statistical (correlations, trends, modes, averages, scatters, and other global or collective properties stated): "Hardwoods tend to occupy the best sites in northern forests."
5. Bounded universal (general fact with limiting qualification): "Hardwoods succeed conifers on the highest sites."
6. Unbounded universal (law-like, applying in all instances of a certain type, in all places, at all times): "Hardwoods succeed conifers."

Example propositions from forest entomology, forest pathology, and silviculture are:

1. "...the greater the diversification of tree species the less frequent will be insect outbreaks" (Graham 1963).
2. "...there is no evidence that the virulence of the [chestnut] blight fungus has decreased" (Hepting 1974).
3. "...tree class 1 achieved a higher increment in relation to basal area than the rest of the stand" (Assmann 1970).

Proposition 2) contains two formal concepts (evidence and decreased), and three factual concepts (American chestnut, virulence, blight fungus). Structures of the factual concepts are analyzed as follows:

> "American chestnut" is first degree (the form is xC) and first order (x is an individual tree). It has qualitative metrical character.
>
> "Virulence" is second degree (the form is hostVparasite) and second order (implied is a large decrease in host population for a small population of parasites). Its metrical character is qualitative, although it comes close to being quantitative because it denotes the intensity of an interaction.

> "Blight fungus" is first degree (the form is yB) and first order. It has qualitative metrical character.
>
> The range of proposition 2) is bounded universal, because the statement is about American chestnut in North America following introduction of the blight fungus.

I leave the identification of factual and formal concepts, their classification, and assignment of predicate structures in propositions 1) and 3) as a reader exercise.

Scientists should be alert to differences in propositions that state their scientific hypotheses. Does the proposition contain a single concept with first degree and first order predicate structure? Or does it contain two or three higher degree concepts with second or third order predicate structure? Concepts with higher degree and, especially, higher order predicate structure, can be immensely helpful in simply stating complex ideas. Propositions alone are not the subject of experimental tests — it is the logical union of proposition, auxiliary assumptions, and initial conditions that is tested. Testability of the union is enhanced if a minimum number of "variables" state the proposition in an algebraically simple manner.

Theory: Like many other words in science, 'theory' has at least two meanings. One deals with an idea and is often used in expressions like "I have a theory on that...." The term 'hypothesis' is more appropriate in such instances.

The appropriate use of 'theory' is when it takes several propositions to state the workings of a thing or system. The union may constitute a theory of how that thing or system operates. However, the union may not. Bunge (1974) calls such a group of propositions a context, and represents it:

$$C = \langle S,P,D \rangle,$$

where S designates a set of propositions,
P designates a set of concepts contained in the propositions,
D designates the domain of the context, and
$\langle$ and $\rangle$ are tuple delimiters.

A theory differs from a context because every proposition in a theory is either a premise or a deduction. That is, the propositions are related by entailment. Theory is thus represented:

$$C = \langle S,P,D,\vdash \rangle$$

where S, P, and D are as above, and $\vdash$ denotes logical entailment (Bunge 1974).

A primary function of theories is to represent more complex systems or processes than can be adequately represented with a single proposition. Theories consist of three categories of statements:

- assumptions,
- premises (definitions and axioms), and
- consequences (theorems and corollaries).

Many scientific ideas require the scientist to make some assumptions about how nature works. Assumptions should be stated explicitly, not left for the reader to guess. Premises consist of definitions and axioms. Definitions are nearly always required to clarify certain ideas, or at least to prevent misinterpretation. They should be given early when stating a theory. Axioms are those propositions believed true about reality — they give the factual science content to the knowledge claims contained in a theory. Each axiom should have a reasonably large domain of "truth" because the domain of truth of a union of axioms may be expected to be less than each proposition individually.

Consequences, the third category of statements in theories, consist of theorems and corollaries. Theorems state deductions that follow from using two or more axioms and as many definitions as needed to anticipate, say, the workings of nature. A theorem is a statement of what nature should be like if two or more of the axioms are treated jointly and are true. A corollary is a subidea of a theorem and asserts a variant of the main consequence asserted by the theorem. Theorems and corollaries are tested in hypothetico-deductive scientific research.

Research planning

Cogency could be increased in research planning if, for example, Gowin's Vee heuristic (Figure 3) were a required part of study plans or study plan summaries (Novak 1979, Novak and Gowin 1984, Stock 1985).

Novak explains: "At the base of Gowin's V[ee] are objects or events that occur in the natural world or that are made to happen .... On the left side... are concepts, conceptual systems, and theories that humans invent.... On the right side of Gowin's V[ee] are the methodological elements of knowledge-making." Implied is a temporal progression up each leg of the Vee from the event.

As developed in Figure 3, the Vee is oriented to discovery research. It may be modified to accommodate justification research.[1] In such reseach one begins with a knowledge claim in the form of a scientific hypothesis — and, following the hypothetico-deductive method — deduces consequences (this completes the left side) which are then tested using controlled experiments and other analysis

[1] Available from the author.

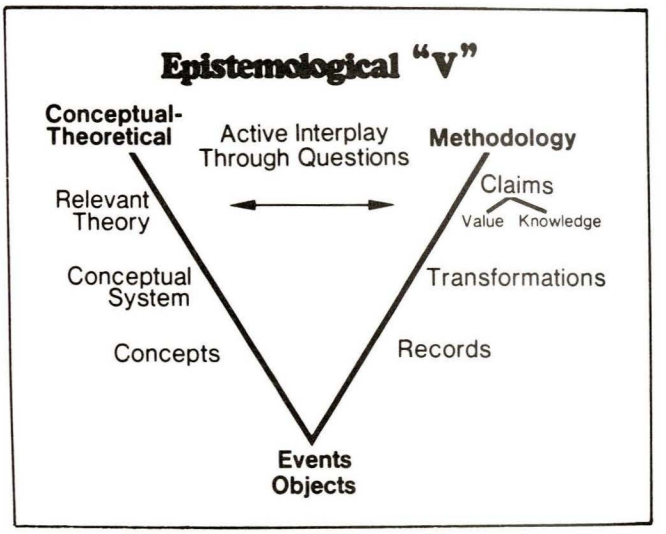


Figure 3. Vee heuristic showing three parts of what Gowin calls knowledge-making: the event of interest [center], constructs proposed to describe or explain the event [left], and methods used to empirically test ideas originating from the left-hand side [right]. [Reprinted by permission: The American Biology Teacher Vol. 41(8), 1979. The National Association of Biology Teachers, Reston, Va.]

methods outlined on the right side.

Whether used for discovery or justification research, the primary value of the Vee heuristic is that it forces a scientist to cover all the bases — a problem statement, ideas about how the problem can be resolved, especially the scientific constructs (concepts, propositions, theories), and the evaluation methods employed to test the ideas against nature.

Research reporting

When reporting research it is not necessary, of course, that every paper contain a detailed analysis of every construct employed. However, it is important that the propositions asserted, theories suggested, and scientific hypotheses offered be clearly stated so the reader is not forced to guess the idea being tested.

Perhaps the quickest way cogency could be increased would be if scientific journals set minimum requirements for manuscript submissions. Most scientific forestry journal articles, e.g., those published in Forest Science, have the format: Introduction, Materials and Methods, Results, Discussion, Conclusions, Literature Cited, essentially identical to that suggested in Day (1979).

Novak (1979) noted that published articles generally ignore the left side of the Vee. To see if Novak's observation applies to forest research publications, I studied 82 articles published in Forest Science Volume 35, looking for statements of ideas the authors were testing, keying on claims about nature the authors offered, or ideas about how nature could be studied more effectively. I counted implied claims at the beginning of a paper or a conjecture formed as the primary finding and expressed at the end.[2] I judged 14 of 82 papers (17 percent) to either begin with, or end by, stating a scientific hypothesis or conjecture of some sort. If only 17 percent of papers report tests or formation of new ideas, what do the remaining 83 percent contain?

DO PERSONAL STYLES OF SCIENTISTS INFLUENCE COGENCY?

Much of doing scientific research in forestry is action-filled — sampling, measuring, calculating, weighing, and mixing. Developing cogency is, on the other hand, very much a cognitive activity, hence it may lack appeal to action-oriented forest scientists.

Gowin's Vee heuristic is helpful in what is called knowledge-making (Novak and Gowin 1984). Part of knowledge-making is discovering and justifying regularities in nature having considerable generality, i.e., laws of nature. In a superb discussion of how new laws are discovered, Feynman (1965) states:

> "... In general we look for a new law by the following process. First we guess it. Then we compute the consequences of the guess to see what would be implied if this law that we guessed is right. Then we compare the result of the computation to nature, with experiment or experience, compare it directly with observation, to see if it works. If it disagrees with experiment it is wrong. In that simple statement is the key to science. It does not make any difference how beautiful your guess is. It does not make any difference how smart you are, who made the guess, or what his name is — if it disagrees with experiment it is wrong. That is all there is to it. ..."

Based on my check of Forest Science articles, I concluded forest researchers are not much inclined toward guessing. In this regard, I am reminded of a quote from Bunge (1967): "Audacity in conjecturing; cautiousness in testing." Of course, neither Feynman nor Bunge is speaking of wild guessing. Guesses must be grounded in what is thought to be "proven" knowledge to qualify as scientific guesses.

[2]Discovery phase research often begins with observational data, but should end with a conjecture. Justification phase research begins with a scientific hypothesis and ends with an evaluation of the hypothesis (McRoberts 1989).

Can one infer anything about the personal qualities of forest researchers from the fact that about 83 percent of published articles are missing even very modestly stated guesses, and nothing approaching audacity? Are forest researchers timid? Or has timidity been taught them as "the method of science"?

ADVANTAGES OF COGENCY

There are several advantages to cogency in scientific research. First, a cogently expressed proposition or theory is more easily learned than a noncogent one. How much of our intellectual heritage are we effectively and efficiently handing down to students of forestry? Scientific reports in North American forestry journals have increased at a 10.8 percent rate during the past 19 years, yet most universities still only require a 4-year program of study. Lewis (1980, 1982) emphasized this point in his attempt to axiomatize 500 important theories of biology. Second, a bit of cogency may bring about more rapid growth of knowledge. By analyzing the concepts behind variables in mathematical equations, scientists may understand why inclusion of one variable in a proposition does a better job of prediction than inclusion of another. Perhaps the better variable designates a concept having higher degree and order structure. Cogent scientific ideas are an excellent source of research projects because most dimensions of the knowledge claim, as well as auxiliary assumptions and initial conditions, will be clearly visible and therefore candidates for further testing. Third, when anomalies arise in proposition or theory application, as they always seem to, scientists will have a better sense of what needs to be "fixed" if either has been cogently expressed.

LITERATURE CITED

Assmann, E. 1970. The principles of forest yield study. Sabine H. Gardner, English translation. Pergamon Press, Oxford. 506 p.

Bunge, M. 1967. Scientific research I: The search for system. Springer-Verlag, New York. 536 p.

Bunge, M. 1968. The maturation of science. [pages 120–147 IN] Problems in the Philosophy of Science. I. Lakatos and A. Musgrave, editors. North-Holland Publishing Company, Amsterdam, The Netherlands. 448 p.

Bunge, M. 1974. Semantics I: Sense and reference. D. Reidel, Boston. 185 p.

Cherrett, J. (ed) 1989. Ecological concepts: The contribution of ecology to an understanding of the natural world. Blackwell Scientific Publications, Oxford, United Kingdom. 385 p.

Day, R.A. 1979. How to write and publish a scientific paper. ISI Press, Philadelphia, PA. 160 p.

Feynman, R. 1965. The character of physical law. MIT Press, Cambridge, MA. 173 p.

Graham, K. 1963. Concepts of forest entomology. Reinhold Publishing Corp. New York. 388 p.

Hepting, G. H. 1974. Death of the American chestnut. Journal of Forest History 18(3):60–67.

Leary, R.A. 1985a. A framework for assessing and rewarding a scientist's research productivity. Scientometrics 7:29–38.

Leary, R.A. 1985b. Interaction theory in forest ecology and management. Martinus Nijhoff/Dr W. Junk Publishers. Dordrecht. 219 p.

Leary, R.A. 1988. Demarcation and distinction rules for scientists. Unpublished manuscript. 14 p.

Leary, R.A. 1989. The role of growth models in forest management decisions. [pages 1–12 IN] Artificial Intelligence and Growth Models for Forest Management Decisions. H. Burkhart, H.M. Rauscher, and K. Johann, Editors. Publication FWS-1–89. School of Forestry and Wildlife Resources, Virginia Polytechnic Institute and State University, Blacksburg, VA. 419 p.

Lewis, R. 1980. Evolution: a system of theories. Perspectives in Biology and Medicine. 23(4):551-572.

Lewis, R. 1982. Theories, structure, teaching, and learning. Bioscience 32(9): 734–737.

McRoberts, R. 1989. A survey of research strategies. [pages 25–31 IN] Discovering new knowledge about trees and forests. USDA Forest Service, General Technical Report NC-135. 114 p.

Novak, J. 1979. The application of phychology and philosophy to the improvement of biology teaching. The American Biology Teacher 41(8):466–470,474.

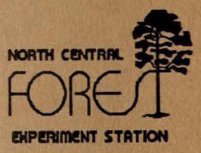
Novak, J. and D.R. Gowin 1984. Learning how to learn. Cambridge University Press. London. 199 p.

Stock, M. 1985. A practical guide to graduate research. McGraw-Hill Book Company, New York. 168 p.

☆ U.S. Government Printing Office: 1991—557-730/40281

U.S. Department of Agriculture, Forest Service.
1991. **Quantity and quality in forest research.** In: Leary, Rolfe A., comp. Invited papers delivered at the 19th world congress of the International Union of Forest Research Organizations; 1990 August 5-11; Montreal, Canada. Gen. Tech. Rep. NC-148. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Forest Experiment Station. 57 p.

Five papers are included that address aspects of quality in forest research. Topics cover (1) the need for a logic of inquiry, illustrated by examining forest fire policy in Western North America, (2) alternative criteria for judging the simplicity of forest models, (3) the identification and analysis of problems and the importance that scientists develop good problem solving skills, (4) a modeling strategy that unites opposites, and (5) the importance of cogency in forest research.

KEY WORDS: Logic of inquiry, simplicity, problem solving, models, cogency.

Our job at the North Central Forest Experiment Station is discovering and creating new knowledge and technology in the field of natural resources and conveying this information to the people who can use it. As a new generation of forests emerges in our region, managers are confronted with two unique challenges: (1) Dealing with the great diversity in composition, quality, and ownership of the forests, and (2) Reconciling the conflicting demands of the people who use them. Helping the forest manager meet these challenges while protecting the environment is what research at North Central is all about.